পঞ্জীভুক্ত নম্বৰ - ৭৬৮ /৯৭ Registered No.-768/97

অসম ৰাজপত্ৰ

THE ASSAM GAZETTE

অসাধাৰণ

EXTRAORDINARY

প্ৰাপ্ত কৰ্ত্তৃত্বৰ দ্বাৰা প্ৰকাশিত

PUBLISHED BY THE AUTHORITY

নং 204 দিশপুৰ, শনিবাৰ, 29 এপ্ৰিল, 2023, 9 ব'হাগ 1945 (শক)

No. 204 Dispur, Saturday, 29th April, 2023, 9th Vaisakha, 1945 (S. E.)

GOVERNMENT OF ASSAM
ORDERS BY THE GOVERNOR
INFORMATION TECHNOLOGY DEPARTMENT

NOTIFICATION

The 26th April, 2023

No. E 272131.- The Governor of Assam is pleased to notify the "The Operational Guidelines for implementation of Assam State Data Policy (ASDP), 2022" which will come into effect from the date of publication in the Official Gazette. The Government also reserves the right to make any amendment to the Guidelines from time to time as deemed fit and proper.

1 Introduction

In an evolving and fast paced digital landscape, the value potential of government data can be realised only through holistic data management guided by standardised frameworks. As a step towards this direction, sequential actions are required towards creation of an interoperable and connected data landscape ensuring security and data privacy, accruing more efficiency, usability, and value-creating opportunities for the State. Under the aegis of the Assam State Data Policy 2022, Government of Assam intends to make a pioneering move to unlock this data potential.

In simple terms, an interoperable and connected data landscape will make digital transformation a possibility. Data collected by any government entity will be available where needed, where security and privacy will be safeguarded, and where adequate legal, technical, and organisational measures will prevent the misuse of data. At the core of data management and governance will be the perspective of approaching data through the lens of individual empowerment, economic growth & recovery, and creating a competitive data democracy. This will require an overhaul of government processes through agile frameworks or iterative implementation to build the momentum for change. The detailed Operational Guidelines under the aegis of Assam State Data Policy 2022 will effectively communicate across the departments to sustainably streamline data management processes through a stream of use cases.

The adoption of the detailed Operational Guidelines has been done after careful consideration of national and global best practices including all the policies, regulations and frameworks currently in place at the central level to make it a competitive reference tool on how key data wealth of an entity is collected, processed, stored, maintained and governed. The Information Technology Act 2000, IT Rules 2011, National Data Sharing and Accessibility Policy 2012, Data Empowerment and Protection Architecture 2020, Data Accessibility and Use Policy 2022, National e-Governance Plan, MeitY e-Governance Standards & Guidelines, National Data Governance Framework Policy 2022, Vision Document of National Data and Analytics Portal, and the Draft Digital Personal Data Protection Bill 2022 has provided the basis for developing this guideline incorporating Assam's operational context.

The Operational Guidelines detail out the measures in the following key dimensions of data management and data governance:

a) *Data Classification:* As digitisation leads to increasing generation of different data types including sensitive personal information, data classification is essential for preserving confidentiality of individuals through authorised restrictions on information access and disclosure, including means for protecting personal privacy and proprietary information; maintaining integrity of datasets by guarding against improper modification or destruction diluting authenticity; and, ensuring availability of datasets for timely and reliable access to and use of information.

b) *Data Acquisition/Collection & Ownership:* Data collection, irrespective of the primary data acquisition mechanisms, shall be ethical and compliant with prevailing data privacy and security legislations and rules thereof. The ownership of data shall reside

with the Data Principals with fair use rights of departments /institutions/ autonomous bodies for which the data was primarily generated.

c) *Lawfulness of Processing of data:* Ensure that data is processed lawfully, fairly, and transparently with appropriate provisioning for special categories of personal data. No personal data shall be processed except for clear, specific and lawful purposes.

d) *Data Interoperability:* Data interoperability will be achieved by way of standardising data elements that appear across datasets, and making data available through an open and machine-readable format, together with their metadata at best level of precision and granularity to help establish common understanding of the meaning or semantics of the data, ensuring correct and proper use and interpretation by its owners and other related users.

e) *Securing the Data System:* Ensure protection of digital information from unauthorised access, corruption, or theft throughout its entire lifecycle.

f) *Grievance Redressal of Data Principals:* Easy and well publicised procedures to make complaints to the Data Fiduciary for expeditious resolution as per the provisions laid down in the Assam State Data Policy 2022.

As the nature of services are becoming increasingly data analysis driven, there is a need for the Government to establish best practices within these dimensions. This will also enable Government agencies to accelerate their data and analytics maturity for effective implementation. As data can be a strategic asset for decision makers, ensuring data and analytics maturity will enable effective governance across departments. As put forth in the Assam State Data Policy 2022, analytics maturity will transform Government of Assam's ability towards harnessing cross-sectoral data generated with public resources to catalyse large-scale social transformation.

The Operational Guidelines have been adopted keeping cognizance of the fact that the *architecture must be technology agnostic* allowing enough flexibility to take into account evolving technologies and standards of compliance. Implementing the Operational Guidelines will also require departments to identify and navigate structural barriers systematically. These efforts will modernise government data infrastructures to unlock significant value across state, economy and society.

2 Objectives and Guiding Principles

The Operational Guidelines are framed with the objective of laying down the principles and direction on data accessibility in both human readable and machine-readable forms while safeguarding citizens' right to privacy, towards enabling data-driven governance with data derivatives serving as public good to enhance government efficiency, improve access to quality public services and delivery of citizen centric benefits, and help advance digital transformation. It shall promote data-usage as a value asset across departments, institutions, and autonomous bodies of Government of Assam, thereby contributing to the overall growth strategy for Assam.

The guiding principles adhered to by the guidelines as per the Assam State Data Policy are given below:

- Openness: Openness provides the foundation for data sharing to unlock its value without compromising on its purpose and utility. For
- Privacy: Privacy is a fundamental right under Article 21 of the Constitution. Privacy of personal data and facts is an essential aspect of the right to privacy; and the Government as the custodian and fiduciary of public data is responsible to safeguard it. Personal data shall be processed or shared only for specific, clear and lawful purposes and in a manner that preserves the privacy of citizens.
- Ethics & Equity: Equity in the access of shareable data while conforming to the highest level of ethical standards.
- Transparency: Transparent mechanisms for public access of open government data for public good with clear traceability to sources of data and information about any intermediate data transformations.
- Legal Conformity: Conform to all laws of the land including the laws enacted by the Parliament and State Legislature on privacy, data security and information protection. At the same time, there will be endeavours to mitigate/remove irrelevant legal barriers on the use of data by citizens and institutions for public good.
- Protection of Intellectual Property: Respect the intellectual property rights of the legitimate data creators/owners by restricting access to IPR protected data. And encourage/advocate increasing use of Creative Commons (CC) or similar public copyright licences, enabling non-commercial uses of copyright protected data-sets to build upon the work of data creator/owner.
- Interoperability & Standards: Enable discoverability and efficient use of existing data towards avoiding duplication and redundancies and establishing techno-administrative protocols for more efficient data flow and usage between data systems.
- Data Quality & Usability: Data being collected, processed and maintained by various entities must be in meaningful and usable format; and is free from anomalies which would inhibit further processing without significant investment on data cleaning and transformation.
- Data Security: Ensure that data and its supporting infrastructure are safe, secure and resilient in the face of established, new and emerging cyber risks.
- Accountability and Formal Responsibility: Ensure accountability and responsibility with respect to the sharing of open data, as well as the adherence of rules relating to access of non-open data.
- Sustainability: Sustainability on all fronts that include technical, economic, financial, legal and other relevant criteria.

3 Definitions

I. **Data:** Representation of information, numerical compilations and observations, documents, facts, maps, images, charts, tables, reports and figures, concepts in digital and/or analog form. It covers all aspects of government functioning including G2G, G2B, G2C.

II. **Data-set:** Collection of logically related features, attributes or variables including processed data or information.

III. **Data Archive:** The digital location where machine-readable data is stored, worked upon / analysed, documented prior to a cut-off past date.

IV. **Data Generation:** Initial collection of data or subsequent addition of data to the same specification. This may be data specifically collected for a particular objective or may be a consequence of the authorised administrative processes of the Government.

V. **Data Principal:** A natural person who is subject to the handled data by which that person can be identified.

VI. **Data Fiduciary/Custodian:** Any person, including the State, a company, any juristic entity or any individual who alone or in conjunction with others determines the purpose and means of processing of personal data;

VII. **Data Processor:** Any person, including the State, a company, any juristic entity or any individual, who processes personal data on behalf of a Data Fiduciary;

VIII. **Metadata:** *Data about data*. The information that describes the data source and the time, place, and conditions under which the data were generated. Metadata informs the user of who, when, what, where, why, and how data were generated. Metadata allows the data to be traced to a known origin and known quality. Metadata consists also of structural aspects such as defining the data and datasets, administrative aspects, such as the processing and audit trail information, descriptive aspects, such as time series and statistical data features (i.e. data source, and the time, place, and conditions under which the data was created) as well as the methods, procedures, concepts, variables, classification, and nomenclature used, including publication date and data coverage.

IX. **Data Standards:** Frameworks that define and embed data handling functions (e.g. data collection, management, transfer, integration, publication); and operate on data in a manner that complies with data format and data syntax specifications produced and maintained by standards bodies.

X. **Information:** Data embellished with a context, in other words, “Processed data.”

XI. **Personally Identifiable Information:** Data about or relating to a Data Principal who is directly or indirectly identifiable, whether online or offline, or any combination of such features with any other information; and shall include any inference drawn from such data for the purpose of profiling.

XII. **Sensitive personal data:** Such personal information which consists of information about Data Principal relating to Password, Biometric information, Official identifier, Financial Data, Information received by body corporate for processing lawful contract or otherwise, Health Data, Genetic Data, Transgender /Intersex status, Health Data, Genetic Data, Transgender /Intersex status, Sexual Orientation and Sex life et. al.

XIII. **Anonymization:** In relation to personal data, refers to the irreversible processes of transforming or converting personal data to a form through which a Data Principal cannot be identified even if the information is combined with other

information, after reasonably considering factors such as time, cost and technology.

XIV. **Aggregation:** Refers to the process of creating higher level data by combining data across Data Principals so that it does not reveal any personally identifying information about a Data Principals. Aggregation is a way of anonymizing data.

XV. **De-identification or pseudo anonymization:** Means the process by which identifiers from personal data may be removed, or masked, or replaced with such other fictitious name or code that is unique to an individual but does not, on its own, directly identify the Data Principal.

4 Scope and Applicability

The Assam State Data Policy lays down the norms and guidelines for collection, collation and processing of data in machine readable format; classification and publishing of open data; secure and restricted access of non-open data across departments; and secure permissioned access of anonymized /de-identified datasets to entities outside of government for bona fide research and analytical studies. These norms are applicable to all data and information created, generated, collected, and archived by departments, institutions, organizations, autonomous bodies of Government of Assam, using public funds provided by the State or the Central Government. The Assam State Data Policy will also apply to data that is recurring in nature and generated owing to automation of government processes and the results emerging out of these for delivery of services and benefits to citizens and businesses, as well as the legacy government data that is still available in non-machine - readable form within the State of Assam.

5 Data Classification

ASDP classifies data into three categories, Open Access Data, Permissioned Access Data and Non-shareable Data.

5.1 Open Access Data

Open Access data are the datasets which are open to use, re-use or redistribution, provided that at most to measures that preserve provenance and openness. The datasets which are published as open datasets would be governed under creative commons license. The data generated by various information systems under the institutional structure of the government will be made openly accessible subject to their compliance with legal and technical grounds of open data classification. The various government datasets would be processed for making certain changes in the structure and content, before these can be notified as open access data. The general guidelines for data anonymization/de-identification would involve, but not limited to, use of the following techniques.

- *Data masking* which involves disclosure of data with modified values. Data anonymization is done by creating a mirror image of a dataset and implementing alteration strategies, such as character shuffling, encryption, term, or character substitution. For example, a value character may be replaced by a symbol such as “*” or “x.”

- *Pseudonymization* is a data de-identification tool that substitutes private identifiers with false identifiers or pseudonyms, such as swapping the "*Jyotirmoy Baruah*" identifier with the "*John Doe*" identifier. It maintains statistical precision and data confidentiality, allowing changed data to be used for creation, training, testing, and analysis, while at the same time maintaining data privacy.
- *Generalization* which involves excluding some data purposely to make it less identifiable. Data may be modified into a series of ranges or a large region with reasonable boundaries. For example, the house number at an address may be deleted, but make sure the name of the lane does not get deleted.
- *Data swapping* – also known as permutation and shuffling – rearranges dataset attribute values so that they do not fit the original information. Switching attributes (columns) that include recognizable values, such as date of birth, can make a huge impact on anonymization.
- *Data perturbation* modifies the initial dataset marginally by applying round-numbering methods and adding random noise. The set of values must be proportional to the disturbance. A small base can contribute to poor anonymization, while a broad base can reduce a dataset's utility. For example, a base of 5 should be used for rounding values like age or house number.

Besides, anonymization of the data, there are various techniques associated with Data aggregation that might be used to the datasets which are (de)classified as open datasets before publishing them in the open data portal. Data aggregation may be done, but not limited to, using the below techniques,

- *In-network Aggregation Approach*
- *Tree-based Approach*
- *Cluster-based Approach*
- *Multi-path Approach*

The above are some of the various techniques which are available in the various Commercial Off the Shelf (COTS) as well as bespoke software products. The various organizations may do the necessary due diligence before adopting these products for impact and performance.

The open access data would be made available easily in machine readable format without any process of pre-authorization, however, the state would keep an automated track of all data seeking request (both intermediary and end-user). Some of the modes of making open access data available include,

- *Open API:* All the open access datasets would be available using OpenAPI Specification [1]using a set of RESTful APIs[2]. Here each of the API must provide an OpenAPI document which describes the API as a part of the dataset publication in the portal. This will allow the third party consuming the API to understand the API structure. The security scheme in the specification should be made optional, so that anyone can access the API for accessing the open access datasets.

[1] The OpenAPI Specification, previously known as the Swagger Specification, is a specification for a machine-readable interface definition language for describing, producing, consuming and visualizing RESTful web services

[2] RESTful API is an interface that two computer systems use to exchange information securely over the internet.

- *Published Datasets*: The dataset which are open in nature may also be available as a complete dataset across the various open data platforms by Government of India as well as Government of Assam. These datasets may be downloaded from these platforms and/or used within these platforms using features like merging, visualization, etc to perform cross-functional analyses.
- *XLS, CSV, etc.* The open access datasets would also be published in various other spreadsheet formats such as xls, csv, etc. The files should not be password protected, and could be downloaded from the portal easily.

5.1.1 **Accessing Open access data:**

The Government datasets which have been (de)classified as open would be published to be accessed using the below portals. The data consumer apps which are accessing open access data related to the State may adhere to the API specifications, as per the Open API policy of MeitY for accessing data published at any of the Central/State Open Government Data portals. Below are the two existing OGD platforms for publishing Open Access Data in compliance to NDSAP 2012[3] and any other legislation by the Union Government governing the publication of open datasets:

- *State OGD Platform* The datasets which are (de)classified as open data would be published at State OGD platform. The data published would be available for consumption in various formats such as xls, pdf, etc as well as available data streams using the pull APIs. The datasets qualified as open will be available for consumption by the third party applications only for the active period as per the dataset ageing given by the Data Provider as part of the Data Exchange Agreement as well as any other guidelines notified by the CDM and Office of the Chief Data Officer (CDO). The open datasets beyond the active period will be available only on request from the open data archive created by the State.
- *Open Government Data (OGD) Platform (data.gov.in):* The state departments, agencies and body corporate may also publish the datasets declassified as open across the OGD platform by Government of India.

5.1.2 **Publication of Open Access Datasets**

The technical agency under the IT Department would be responsible for the final upload of the datasets on to the State OGD Platform, however, the Seniormost secretary of the Data Provider department would be the final authority for publication of the Open Access Datasets. The CDO would be the facilitating the publication by guiding the Data provider department, and also auditing for the compliance to Assam State Data Policy 2022 for publication of the Open access datasets in the State OGD Platform as well as the central OGD platform.

Compliance Grid for publication of Datasets

Below is the step-by-step approach for publication of datasets as Open data:

[3] National Data Sharing and Accessibility Policy-2012

Steps to be taken	Creator	Checker	Approver
Step 1: Departmental Data officer to upload datasets on the State Open Government Data Platform for review and publication by the Seniormost Secretary of the Department in compliance to ASDP 2022.	Departmental Data/IT team	Departmental Data Officer (Nodal Officer for Smart Governance)	Seniormost Secretary of the Department.
Step 2: Technical agency under the IT Department maintaining the portal to review submitted datasets in the Staging Area for coherence with Open Data Publication Metadata Standards, Attribution Template Standards, Usage License Clarity and forward to Seniormost Secretary of the Department for approval	Technical agency	Seniormost Secretary of the Department.	N/A
Step 3: Seniormost Secretary of the Department to approve publication of dataset or send them back for revision	Seniormost Secretary of the Department	N/A	N/A

Review and audit

The office of the CDO will be doing periodic audits of the datasets which have been published as Open datasets in the State OGD and Central OGD portals. Any non-compliances would be communicated to the Seniormost Secretary of the Departments which has published the dataset.

5.1.3 **Processing of the Open datasets:**

The third party applications accessing the open datasets may process the data for non-commercial and/or research/educational purposes only guided by the Union and State laws. Below are some of the guidelines for usage of open datasets by third parties:

- *Use of Open Data.* The third party apps may retrieve, download, copy, modify, translate, adapt, distribute, sort, search, and reuse Open Data for any lawful purpose only. The third party apps shall not reproduce any trademark or service mark on the Open Data Portal/website or in Open Data without the express written permission of the respective Data owner. Any unauthorized attempts to upload information to, modify, or cause damage to the State Data Analytics Portal/website is prohibited.

- *Indemnification.* The third party may have to agree to indemnify, defend, and hold harmless the State, and the State's respective officials, agencies, officers, departments, autonomous bodies, employees, licensors and agents, from and against any and all liabilities, loss, claims, damages, costs and/or actions (including attorneys' fees) based upon or arising out of any breach by them of

their obligations including, but not limited to, any claim or cause of action, of any type, that may be or is alleged to have arisen out of use of the Open Data from State. Notwithstanding the indemnification obligation, the Centre for Data Management (CDM) would reserve the right as its sole discretion to defend any such claim and would agree to provide the Data Owner with such reasonable cooperation and information as requested by him/her.

5.2 **Permissioned Access Data:**

The various Government of Assam departments, associated institutions / agencies as well as body corporate would collect data through various sources of data acquisition as mentioned in section 6.1.1. One of data acquisition sources could be the transactions emanating from the various services (G2C, G2B, G2G[4]) provided by them, or the various schemes implemented by them. Based on the sector/domain of which the Government of Assam departments, associated institutions / agencies as well as body corporate are part of, they would have generated data which could be classified as data generated at the same source as the institution and is the primary data. This data would be collected via forms implemented at the information systems which caters to the government functioning.

Government of Assam departments, associated institutions / agencies as well as body corporate also acquire data streams from external sources such as social media, IOT[5] sensors, geo spatial data, etc. which could be used for functioning of the department work. This data could be classified as the Big data. Below are the criteria set for Data Owners (Government of Assam departments, associated institutions / agencies as well as body corporate) to designate any specific data set as permissioned access data.

- Datasets which are maintained as electronic registries across the Management Information Systems (MIS), and act as a single source of truth regarding certain actors such as citizens, employees, beneficiaries, etc.

- Datasets which are generated by the Management Information Systems (MIS) of the particular government institution pertaining to transactions emanating from providing public services or running schemes which would be accessible only to that institution by default. The ownership of such generated data would remain with the same government institution, and any organization which wants to access such data would need permission.

- Personal data of the people (employees, beneficiaries, pensionaries, etc.) which are collected by the government institutions are owned by the Data principal, with consent been provided to the Data processor/Fiduciary to be used for providing certain services to him/her.

- Datasets (transactional) which have been acquired from a different government agency (source MIS) for implementing certain schemes or providing composite

[4] G2C: Government to Citizens; G2B: Government to Business and G2G: Government to Government

[5] IOT: Internet of Things.

services with requisite permissions, would still remain as a permissioned dataset when used in the destination MIS.

- Datasets generated from the diagnostics tests, either biochemical or readings generated from the machines will be regarded as Personal Health Information (PHI) of the individuals/groups/families. This data will be regarded as personal information of the Data principal, and would require consent before those can be shared. The provision of deemed consent will not be applicable for PHI of Data Principal; and explicit prior consent will have to sought for the purpose of processing and/or using PHI dataset.

- Datasets which have been collected by means of non-human collection such as via bots, drones, aerial imaging, satellite and other emerging IoT devices by government institutions, and are essential to be maintained either as master data or secondary data for providing services or running schemes as well as generating insights for decision making by the institutions will be either maintained in the MIS or the Central Data Lake for further cross-functional analysis.

Access to Permissioned access data

The datasets across the departments which are generated across the various information systems owned by the respective Government of Assam departments, associated institutions/agencies as well as body corporate will be shared by the departments/agencies delivering composite government services, execution of development schemes which require datasets from other departments for eligibility and other checks as well as performing cross-functional analysis across the various Government of Assam departments, associated institutions/agencies as well as body corporate for process improvement. Access to these datasets would be governed by the rules of access as mentioned in Section 6.11 of this document.

Access to Personal Data

The various Government of Assam departments, associated institutions/agencies as well as body corporate shall demonstrate the grounds for processing of personal data.

- The Data Processor may process the personal data of a Data Principal only in accordance with the provisions of the ASDP 2022 and/or prevailing norms on data privacy and data protection by Union or State legislation, for a lawful purpose.

- On or before using the personal data of the Data Principal, the Data processor/Data fiduciary may seek for his/her consent, a Data Fiduciary may seek the consent from the Data Principal as per rules/guidelines of the existing Union and State laws.

Retention of Permissioned access data

The personal data retention would be guided by the lawfulness of processing of personal data as guided by the ASDP 2022 or the prevailing norms of Data Protection and Privacy to be enacted by the Parliament of India. The datasets which are classified as

permissioned access by the Assam State Data Policy 2022, would be retained as per the departmental data retention guidelines published from time to time by the Data Processor/ Fiduciary / Custodian of the data based on the requirements of external stakeholders such as Auditor General (AG), Judiciary, etc. Big data which would be collected in the Central Datalake would be guided by the data retention policies of Big data notified by the Chief Data Officer (CDO) from time to time.

Localization of Data

Permissioned access data (which includes even the sensitive personal data) generated within India must be retained in the servers within the Indian geography which are also governed by the sectoral laws such as,

- The Reserve Bank of India's Directive 2017-18/153 (April 6, 2018) issued under the Payment and Settlement Systems Act 2007 which is mentioned below:

It is observed that not all system providers store the payments data in India. In order to ensure better monitoring, it is important to have unfettered supervisory access to data stored with these system providers as also with their service providers / intermediaries/ third party vendors and other entities in the payment ecosystem. It has, therefore, been decided that:

i. *All system providers shall ensure that the entire data relating to payment systems operated by them are stored in a system only in India. This data should include the full end-to-end transaction details / information collected / carried / processed as part of the message / payment instruction. For the foreign leg of the transaction, if any, the data can also be stored in the foreign country, if required.*

- Section 128 of The (Indian) Companies Act 2013 (ICA) as mentioned below.

(1) Every company shall prepare and keep at its registered office books of account and other relevant books and papers and financial statement for every financial year which give a true and fair view of the state of the affairs of the company, including that of its branch office or offices, if any, and explain the transactions effected both at the registered office and its branches and such books shall be kept on accrual basis and according to the double entry system of accounting:

Provided that all or any of the books of account aforesaid and other relevant papers may be kept at such other place in India as the Board of Directors may decide and where such a decision is taken, the company shall, within seven days thereof, file with the Registrar a notice in writing giving the full address of that other place:

- Paragraph 3(9) of The IRDAI (Maintenance of Insurance Records) Regulation, 2015.

3. Maintenance of Policy and Claim records

(9) The records including those held in electronic mode, pertaining to all the policies issued and the claims made in India shall be held in data centres located and maintained in India only.

Maturity of Data Systems (Information Systems)

The various data systems which will be holding the permissioned access data need to be measured for maturity, so that the scope for improvement can be identified. An index known as **State Digital Transformation Index (SDTI)** will be developed which will allow ranking of the Government of Assam departments, associated institutions / agencies as well as body corporate on the basis of maturity of Data Systems (Information Systems).

The proposed approach of constructing the index has been detailed in the Appendix section of this document.

5.3 Non –shareable/Sensitive data

As per the ASDP 2022, sensitive personal data and the datasets which are confidential in nature and are in the interest of the country's security in not opening to the public would be considered as non-shareable data and would fall in the negative list. This includes data/information that is expressly prohibited from disclosure as per exemptions defined under sections 8 and 9 of the Right to Information Act, 2005, the Aadhaar (Targeted Delivery of Financial and Other Subsidies, Benefits and Services) Act, 2016, the Collection of Statistics Act, 2008 and rules framed thereof; also the provisions of the Digital Personal Data Protection Law, once enacted.

Initial list of Negative list of non-shareable data

Rule 3 of the Information Technology Act (IT) Act rules 2011 list down the following eight types of data (personal) as sensitive data. This list will be applicable for Government of Assam departments, associated institutions / agencies as well.

- password;
- financial information such as bank account or credit card or debit card or other payment instrument details;
- physical, physiological and mental health condition;
- sexual orientation;
- medical records and history;
- biometric information;
- any detail relating to the above categories as provided to body corporate for providing service
- any of the information received under above categories by body corporate for processing, stored or processed under lawful contract or otherwise

Besides the above, several sections under IPC for crime against women as well as Indian Witness Protection Scheme, 2018 seeks to protect the information of the victims and witness as they consider the data of sensitive nature.

The Office of the CDO and the Centre for Data Management (CDM) will be reviewing the list of non-shareable data from time to time based on the provisions of Union law on Personal Data Protection. Non-shareable data may be declassified as permissioned access data or open access data. Below are the steps for Notification of the negative lists across the state.

Steps to be taken	Creator	Checker	Approver
Step 1: Review Departmental Negative lists for coherence with personal/sensitive personal data identifiers mentioned	Departmental Data/IT team	Departmental Data Officer (Nodal Officer for Smart Governance)	N/A

Step 2: Forward Negative Lists to the Seniormost Secretary of the Department for approval	Departmental Data Officer (Nodal Officer for Smart Governance)	Seniormost Secretary of the Department	Seniormost Secretary of the Department
Step 3: Notify Negative Lists as official negative list for the Department of Government of Assam	Seniormost Secretary of the Department	N/A	N/A

Declassification of non-shareable data

Non-shareable data or personal data can be shared by the Department(s)/Government Organization(s) after declassification based on legitimate grounds of sharing, will be governed by the following protocols:

- Data sharing as permitted/mandated by any extant law, regulation, ordinance, order, bye-law, rule or notification of the Government of India or the Government of Assam, order of any competent judicial or quasi-judicial authority, carrying out the obligations under employment, social security or social protection law, or a collective agreement between the Data principal and other parties.

- Unless explicitly prohibited by law, sensitive data may be shared as permitted by the Data Principal on the basis of explicit Consent, if any, as per the prevailing norms on Data Privacy and protection across the nation.

- If necessary to protect the vital interests of the data subject who is physically or legally incapable of giving consent, such as a child or physically or mentally disabled person, as per the prevailing norms on Data Privacy and protection.

- For reasons of public interest in the area of public health, such as protecting against serious cross-border threats to health or ensuring high standards of quality and safety of health care and medicinal products or medical devices.

6 Data Management and Governance Framework

6.1 Data Acquisition and Ownership

6.1.1 Data Acquisition Methods

Data would be acquired by the Government of Assam departments, associated institutions / agencies as well as body corporate by the below methods:

- *Collection at source:* Most of the data collected through the MIS for delivering public services or delivery of government programmes are collected at the source. Before this data is collected, strict controls of Data Quality assurance must be in place. All of the analysis, definitions, and standards need to be in place prior to any field information collection. Data must be reviewed and updated on a regular schedule to maintain a high standard of quality. Metadata must also be updated at the same

time. Data is also collected through non-human means e.g., data collected from industrial sensors, satellites, drones, etc.

- *Legacy Data*: The data might also be collected from the legacy sources. The legacy sources could be in electronic as well as in non-digitized format. The non-digitized data are usually digitized before it becomes usable. However, a few considerations must be there before using the legacy data which includes ensuring acceptable data quality, cost of acquisition, usable format of the data, and reliable storage medium.
- *Shared/Exchanged*: Data is also acquired from external data sources through data sharing or exchange. The key considerations of acquiring data from external sources include,
  - Data exchange agreements which need to include provisions concerning access and dissemination. This would be governed by Assam State Data Policy 2022.
  - Data organization which refers to availability of the data in usable form. Also, the need to consider the cost associated with conversion/transformation of the data to make it usable. Data which is shared to be archived in Assam State Datalake must be available in the usable form, so that only limited transformation of data is needed to be done by the custodian.
  - Data must have the corresponding metadata and other pertinent documentation such as data dictionary.
  - Data acquired through this means must be considered for completeness. Also, the cost of addressing the gaps in data must also be considered.
- *Purchased*: Data may also be purchased from the agreeing parties which should be governed by a contract/agreement. The data licensing agreements must also be considered to understand the restrictions placed on the use of this data.

6.1.2 **Data Standards**

Data Standards are produced by consensus of the experts and are ratified by a standards authority such as International Standards Organization (ISO). Standards make it easier to create, share, and integrate data by ensuring that the data are represented and interpreted correctly. Standards also reduce the time spent cleaning and translating data. The standards which are generally included are,

- *Dataset-level Standards:* Dataset-level standards, normally documented with a data dictionary, specify the scientific domain, structure, relationships, field labels, and parameter-level standards for the dataset as a whole. Local Government Directory (LGD) (URL: https://lgdirectory.gov.in/) is a good example of. It is mandated that all the Government of Assam departments, agencies as well as body corporate must follow the LGD standards while capturing geographic data related to any entity.

- *Parameter-level Standards:* Parameter-level standards define the format and units for a given parameter or field within a dataset and help users correctly interpret the values. Parameter-level standards should be adopted at the time of data collection, that is when values in a field are created or recorded. If standardization of a parameter in an existing dataset will result in any loss of original detail

or information, a best practice is to retain the original parameter and add a separate field for the standardized parameter. Some examples are given below:

	Event Date	Decimal Latitude	Decimal Longitude	Country code
Value Examples	2010-05-17	42.33	-98.1449	US
Parameter level standard	ISO-8601	ISO 6709:2008	ISO 6709:2008	ISO 3166-1 alpha-2

- *Data encoding and Interface standards*: Most dataset-level standards also offer guidance on how to encode data. Data encoding standards define the rules for structuring and organizing data for use in a given context. These standards ensure that when applications read data, the information and context is preserved. Data encoding standards are generally associated with file formats. Common data encoding schemes include ASCII, Unicode, hexadecimal, Base64, and MIME. Some data encoding systems may also result in data compression, such as gzip.

- *Documenting data standards in metadata:* Parameter-level and dataset-level data standards should be documented in the accompanying data dictionary and metadata record. The details regarding the metadata standard can be found in section 8.6.

6.1.3 Planning for Data Entry (Use of Data templates)

A plan should be developed for each step of the data collection, data entry, and data storage process. Data templates allows these processes to be controlled and standardized, reducing potential for human error. Some of the best practices for templates are given below:

- *Data Templates*
  - Use of electronic forms when entering data so that some quality assurance is automatically taken care of during data entry. Usage of standardized forms across the Information Systems (IS) is recommended.
  - Consistency
    - Be sure that the columns of data have only numbers, dates, or text and contain the same information.
    - Names, codes, and formats must be consistent.
    - Dates and geographic data should use the same format and datum (for geographic data).
  - Efficient data organization - Data that are organized efficiently can be read by statistical packages.
  - Descriptive names
    - Column names should be descriptive and should not contain strange symbols.

        - The file name should be descriptive and concise.
    - Standardize how missing data are represented. Decide how you are going to handle missing/null values before entering data. Be sure if you use a specific value to represent missing data, there is no chance that value will be used to represent actual data at a later time (e.g., using 9999 to represent missing data).
        - Use a separate column to track and describe missing data across the rows.
    - Create forms to control how the data are entered.
        - Only allow certain data types (i.e., numbers, letters) to be entered in designated columns so that if the data type is entered incorrectly, the form will reject the entry.
    - Check the beginning and end portions of gathered data for errors, and then randomly spot-check other values throughout the form.
    - Consider graphing or mapping of collected data to ensure there are no unexpected / unexamined outliers.

- *Long term storage*
    - Use file formats consistently, and preferably formats that will remain readable in the long-term, e.g.,
        - ASCII, UNICODE, non-proprietary, unencrypted, uncompressed
        - Use comma-delimited ASCII files to represent tabular data

6.2 Lawfulness of Processing of data

A person may process the personal data of a Data Principal only in accordance with the prevailing norms of data privacy and protection guided by the Union or the State legislations on this subject.

For the purpose of this section, "lawful purpose" or "lawfulness" means any purpose which is not expressly forbidden by law.

6.3 Protection of privacy and intellectual property rights of the Data Principals

The infringement of privacy of the Data principal arising out the unauthorised access to personal as well as non-shareable data would be protected under the Assam State Data Policy 2022 and/or any other prevailing Union or State Laws.

6.3.1 **Redressing Grievances of Data Principals**

- Right to raise a grievance with Data Fiduciary
    - Data Principal may raise his/her grievance with the Data Processor/Fiduciary. The grievances may be in contravention of any of the provisions of ASDP 2022 or the rules made thereunder, which has caused or is likely to cause harm to the concerned Data Principal.
    - Such complaints may be lodge at the ASDP 2022 helpdesk to be set up by the Centre for Data Management (CDM) which will allow the complainant to contact the helpdesk via email, phone, text, and other means as necessary. The

detailed procedures for raising a grievance would be available with the helpdesk.

- Such complaints /grievances shall be resolved by the Data Processor/Fiduciary in an expeditious manner and not later than thirty days from the date of receipt of the complaint.

6.4 Data Interoperability:

Design considerations for interoperable Information Systems: The State Data policy endeavours to establish an environment for autonomous information systems built on principles of building digital ecosystems such as Aadhaar, UPI, Account aggregator, GSTN, etc. which are interoperable by Design, and are guided by the following principles. These principles are also aligned to InDEA 2.0 framework by MeiTY.

6.4.1 **Federated Architecture:**

The information systems must be built around the constructs of Single-Source-of-Truth and System-of-Records, both of which should be created and maintained by the entity legally responsible to do the same. All other entities shall be required to access it from such a source/ record.

- *Open API-based:* All the Government of Assam departments, associated institutions / agencies as well as body corporate must adopt the principle of 'Open-API by default', exceptions shall be justified. The Open API Policy (https://www.meity.gov.in/writereaddata/files/Open_APIs_19May2015.pdf) and the guidelines issued by the Government of India shall be followed, and the specifications on Open APIs published by global organizations like OAS Foundation (Open API Specification) may be adopted as needed.

- *Electronic registries*: All information systems require master data or actor (person/entity/thing) data related to that system to be maintained for identification, validation, etc. This data must be maintained in electronic registries to be accessible with Open API for other applications to seamlessly validate as well as use attested and authenticated data. This is even more critical when it comes to people and entities where various claims can be electronically validated against such registries via open APIs avoiding paper-based validations, thus increasing trust while decreasing cost of validation. For ex – Aadhaar is the registry of "usual residents of India", PAN system is the registry "people/entities" who are direct taxpayers, PDS database is the registry of "people (and families) who receive food subsidy, and so on. All the registries must be accessible only by the Unique digital ID. Custodians of such registries should ensure appropriate policy is applied to either allow user controlled uniqueness or state controlled uniqueness. In addition, the fields in the record of that subject are to be verified/attested or marked as self-declared. When registering, people must be given an option to use their existing digital IDs such as Aadhaar, mobile, etc as appropriately to fit the purpose of that registry and also allow people to control, update, manage their record using the common IDs such as Aadhaar, mobile, etc.

- *Digital ID*: A digital identifier will be the "key" to a registry where the subject (ID holder) is present who, in turn, is empowered to control his/her ID, manage the registry record (his/her profile in that registry), choose to use it for availing other 3rd party services through authentication and consented eKYC (digitally signed profile sharing). These identifiers may be purely numeric (e.g. Aadhaar number, mobile number, health ID within Ayushman Bharat Digital Mission, etc.) or alpha-numeric (e.g. PAN number, Vehicle number, UPI Address, etc.) with or without any logic attached in generating the identifier itself (random vs logic based identifier).

- *Federated registries*: The registries built stand-alone by various information systems may get interlinked via registry IDs depending on the policies that allow such linking. Electronic registries when linked via the IDs to allow easy, paperless onboarding of citizens and also avoid repeated data verification needs. For example, when a beneficiary is registered for, say, PDS scheme, that record will be linked to Aadhaar by the PDS system storing the Aadhaar number (or a tokenized version of it). When a registry allows users to use "existing IDs from other registries" to be used as an authentication mechanism, it not only creates an "auto verified/attested" set of fields in the new registry (registry provider does not have to re-verify those fields again), but also gives convenience to the people to reuse and leverage commonly used IDs. To achieve this, it is essential that all registries are built to allow single sign-on (SSO) using existing IDs in other registries as well as expose itself as an SSO provider for the next set of systems. *Until the SSO is established between the various departments for accessing departmental registries, the data would be shared between the Data Requestor and Data Provider through the State Data Exchange.*

- *Credentialing:* The government, academic, industry, and other ecosystems today issue many certificates (government and non-government issued), licenses, authorization letters, etc. These information systems may adopt issuance of Virtual Credentialing which could be easily issued, and should adhere to the following design principles.
    - Verifiability: Authenticity of the credential should be digitally verifiable by any application to which it is presented in a paperless and presence-less manner.
    - Portability: To ensure empowerment and choice for the credential holder, the credentials should be digitally portable across systems participating in the ecosystem.
    - Permanence: The credentials should continue to exist and be valid beyond the lifetime of the institution where it was awarded.
    - Self-Describing: The credential model should be self-describing in a manner that the verifier of the credential does not require private sources of information to validate or understand it.
    - Consent-based: Design must ensure privacy preservation across the system including taking holder's consent for collection and use, either directly or indirectly using the Consent Manager.
    - Inclusive: Design of credentials must ensure inclusion in terms of digital and physical usage (through printed modes with signed QR codes etc.),

multi-lingual support, online and offline usage, and work seamlessly with inclusive and accessible technologies.

Various information systems implementing VCs should set up credential issuance platforms to issue standard schema based W3C VC complaint credentials, give choice to users to download it, access via Digilocker[6], and also via any additional channels such as emails, instant messaging platforms, blockchain based systems, etc. Such issuance platforms should provide both natively digital formats along with a printable format which contains a digitally authenticated QR code. In addition, a process for issuance through revocation (as necessary) be implemented in such platforms.

6.4.2 **Cloud Hosting:**

Cloud should be the first choice in taking ICT infrastructure decisions, either hosted at State Data Centre (SDC) of Government of Assam or National Data Centre (NDC) of Central Government, or with any of the MeitY empanelled Cloud service providers. Cloud and virtualization technology offers the support to the application in scaling up in a short period of time. Permissioned Access and Non-shareable Data must always be hosted in NDC/SDC or Government Community Cloud[7] (GCC)/Virtual Private Cloud. In other cases, public or hybrid Cloud may be chosen depending on the requirements of the Government department, associated agencies or body public.

Note: Depending on the nature of the data and capabilities of State Data Centre (SDC), the implementing agencies may consider to host the data at the State Data Centre (SDC) till the compute/storage utilization of the Private Cloud reaches 60% of the net capacity. Post that, the hybrid mode of hosting may be preferred.

6.4.3 **Privacy by-Design:**

The information systems must design and publish system policies which conforms to the principles of privacy-by-design in capturing and storing personal data. This should conform to guidelines of handling personal data as per the State Data policy, as well as prevailing Union and State laws. Any obligation to capture electronic consent from the beneficiary should follow the MeitY guidelines as defined here.

6.4.4 **Security by-Design:**

The implementing agencies must design and enforce a cybersecurity policy that conforms to the principles of Security-by-Design, and an ISMS (Information Security Management System) that conforms to the ISO guidelines related to information security. Below are few indicative principles,

- Minimize attack surface area
- Establish secure defaults
- Follow the principle of least privileges

---

[6] DigiLocker is a flagship initiative of Ministry of Electronics & IT (MeitY) under Digital India programme. DigiLocker aims at 'Digital Empowerment' of citizen by providing access to authentic digital documents to citizen's digital document wallet

[7] Government Community Cloud is one of the services to be provided by the MeitY empanelled Cloud service providers where Government data need to be maintained together at one place.

- Follow the principle of defence-in-depth
- Fail securely
- Don't trust 3rd party services
- Observe principle of separation of duties
- Avoid security-by-obscurity

The information systems must also adhere to the Assam Cyber Security Policy 2020 for guidance on general rules for information security compliance at the State level.

6.4.5 **Non-repudiation and Data provenance**

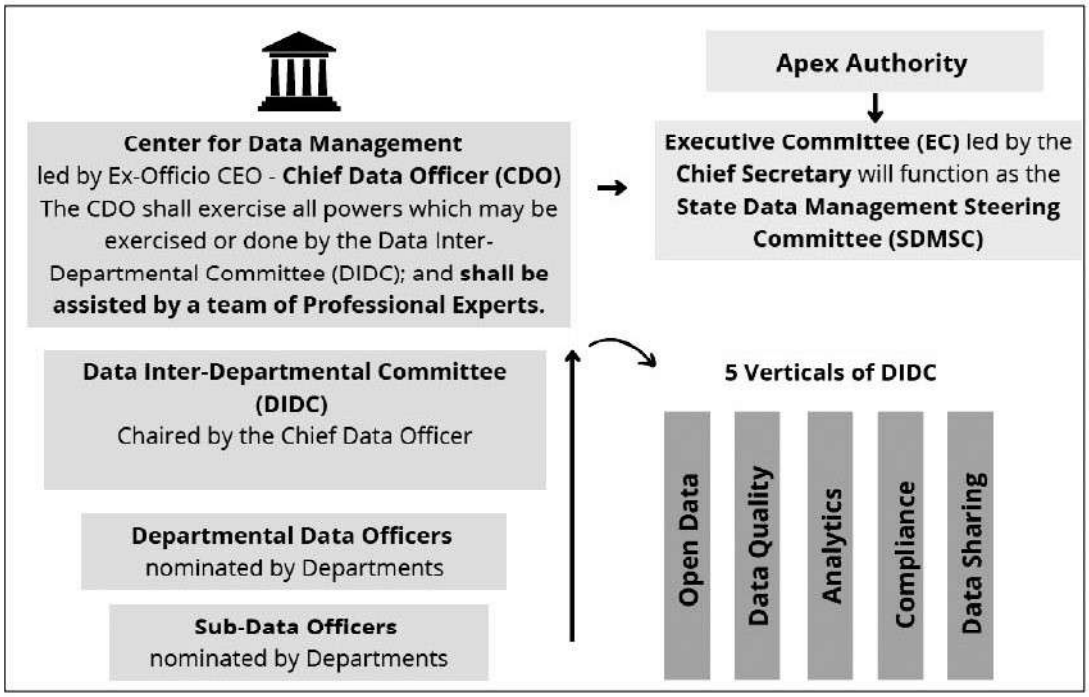
Data and documents either entered or generated from within the departmental information system (IS) should be non-repudiable which refer to the ability to ensure that a party to a contract or a communication cannot deny the authenticity of their signature on the data or a document or the sending of a message that they originated. A non-repudiable audit trail must also exist for all changes made to entries to the departmental registries, and also the transactional data stored. No data shall be deleted from registries, though in certain circumstances, the registry owner may choose to mark certain data objects as obsolete/ passive.

6.4.6 **Differential privacy**

Based on the Data Requests from the Requestor organization, the various datasets which are required to be accessed for planning various schemes or for making the aggregated data available as open data could be made available by Data Provider after implementation of differential privacy algorithms. This enables deidentification of the private information of the individuals in case the consent for sharing the required data does not exist with the Data Processor or Fiduciary.

6.5 **Techno-administrative implementation**:

Article 9 of Assam State Data Policy 2022 has delineated in detail the technical and administrative paradigms and mechanism of implementing the data policy. This has been summarised in the illustration below; wherever necessary the Operational Guideline draws upon this mandate to frame / enunciate / elaborate procedures to ensure effective implementation of ASDP 2022.



The information systems both public and private built using the above principles intended to provide G2C, G2B and G2G services will be architected as part of Open Ecosystem Networks which will be driven by open protocols. The implementation of such an open network will be made possible by adoption of open protocols such as digital creative commons, and the mechanism to manage, co-ordinate, incentivize a state-wide adoption and long-term sustenance of this open network. The Chief Data Officer (CDO) supported by CDM would be responsible for defining the governance mechanism for this open network from time to time, which will be essential to build trust on the network protocol by the Government of Assam departments, associated institutions / agencies as well as body corporate communicating with each other. (Detailed guidelines on governance mechanism of the data exchange is provided at the below sections).

Some of the enablers of the open network protocol for communication and data sharing across ecosystems are cited below:

6.5.1 **Overview and reference architecture of State Data Exchange:**

A mechanism of Data exchange would be set up at the State level to enable controlled exchange of data between the information systems. The State Data Exchange (DE) would be based on the online marketplace model of consumer goods platforms like Amazon, Flipkart, etc., however, this won't be a platform, but an open network protocol which will enable sharing of data across the network participants without storing any data in it. Below are the objectives of the DE framework:

- Enable, encourage, and authorise data exchange to increase the value of the investment in government data and ease the sharing and/or integration of data between Government of Assam departments, associated institutions / agencies as well as body corporate communicating with each other.
- Reduce the cost and resource intensity of data exchange by creating a standardised Government of Assam data exchange approach (regardless of data

type, data exchange method or protocol) and ensure responsibilities for the data after exchange are clear and agreed prior to exchange.

- Retain data integrity by considering the quality, value and authenticity of the data being exchanged.
- Balance the need for safety and transparency in data exchange with the need for better informed decisions, evidence-based policy development, performance reporting and operational efficiency.
- Ensure data exchanges are fit-for-purpose (i.e., meet government needs) and able to be supported.

Note: The data referred in this section refers to 'structured data', and 'unstructured data' will be covered in section on Assam Datalake.

The DE would consist of three major components in its architecture:

- Catalogue services: This server would be analogous to online catalogue services of an online marketplace with the following functions:
  - Search and discovery of data resources
  - Providing description of the data resources available at the DE.
  - Publication of the APIs for various types of search such as text search, geo-spatial search, attribute search and relationship search.
  - CRUD operations on the metadata
- Authorization services: This service would act as a checkpoint for access to permissioned access data and non-shareable data as specified by the Data processor. This service would consist of APIs for operations associated with request, validate and manage an access token for data resources, and policies associated with access to a data resource.
- Resource Server: Provides data access for the data resources available with the DE. The resource services deliver data to data consumers in compliance with the access policy requirements set by the Data processor of the resource. For this compliance, a resource server must implement a token introspection interface with the authorization server.

The State Data Exchange is envisioned to adhere to the Unified Data Exchange Architecture published by MeitY.

The detailed functional and technical design of the State Data Exchange will carried out as a downstream activity before the actual implementation is initiated.

6.6 **Enabling metadata models***:*

CDM would be the nodal agency for specifying the metadata catalogue and associated activities. A comprehensive and mandatory Meta Data Catalogue for all departments will be prepared and updated from time to time which contains the definition of each of

the citizen centric fields and master data, including the source of truth and responsible data processors or data fiduciary.

- Metadata catalogue contains definitions, datatypes, length, and all other attributes of a typical data element like name, address, occupation etc. These kinds of fields are stored in multiple ways in multiple applications.
- Metadata catalogue must be compliant with the standards set as per the State Date Exchange.

CDM may recommend the metadata model in the lines of *Dublin Core Metadata Element Set*. The current standard defines fifteen elements. Each of the elements is described as below:

Dublin Core Element	Use
Title	A name given to the resource.
Subject	The topic of the resource.
Description	An account of the resource.
Creator	An entity primarily responsible for making the resource.
Publisher	An entity responsible for making the resource available.
Contributor	An entity responsible for making contributions to the resource.
Date	A point or period of time associated with an event in the lifecycle of the resource.
Type	The nature or genre of the resource.
Format	The file format, physical medium, or dimensions of the resource.
Identifier	An unambiguous reference to the resource within a given context.
Source	A related resource from which the described resource is derived.
Language	A language of the resource.
Relation	A related resource.
Coverage	The spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant.
Rights	Information about rights held in and over the resource.

It is recommended that all the mandatory fields must be included. CDM may adopt or reject the elements of the metadata based on the implementation considerations, flexibility and adoptability across the state.

6.7 Assam State Datalake as unified source of archived datasets:

Government of Assam departments, associated institutions / agencies as well as body corporate over a period of time would generate big data from the various sources. Big data in government is the influx of data from disparate sources such as traffic and CCTV cameras, sensors, satellites, body cameras, calls, emails, direct messages, and social media, as well as the use of emerging technology from private IT spaces and academia for evidence based policy making for data driven governance of the public sector. Big data captured in government and public sector would be classified as Open access data, permissioned access data and non-shareable data.

To manage the influx of the big data, derive correlation and generate insights, it is important to set up the Assam state Datalake which would store and process both unstructured as well as structured data. The characteristics of the big data in the Datalake must include

- Volume: overall volume of data
- Variety: various variety of data formats
- Velocity: the pace of data being generated
- Value: economic output of data processing
- Veracity: the quality and trust of data processing

These are also the guiding principles of implementation of the Big Datalake. The central Datalake would be an important pillar of cross-functional analysis of data.

Design considerations for Datalake

The Big Datalake must be prepared by adherence to the following design considerations:

- *Data Ingestion*: Datalake must support connectors to get data from different sources and load into the Datalake. The Data ingestion function of Datalake must support:
    - All types of structured, semi-structured and unstructured data.
    - Multiple ingestions like Batch, Real-time, One-time load.
    - Various types of data sources like databases, webservers, emails, RSS feed[8], social media feeds, IoT, and FTP.
- *Discovery*: The design of the Datalake must have a simple organization, so that it is understood by all the stakeholders. All the big data generated by the State government agencies must be available at one place.
  The Data tagging technique could be used to express the data understanding, by organizing and interpreting the data ingested in the Data lake.

[8] RSS feed: RSS is a web feed that allows users and applications to access updates to websites in a standardized, computer readable format.

- *Security:* Security needs to be implemented in every layer of the Data lake. It starts with Storage, Unearthing, and Consumption. The basic need is to stop access for unauthorized users. It should support different tools to access data with easy to navigate GUI and Dashboards.

  Authentication, Accounting, Authorization and Data Protection should be some important features of data lake security.

- *Storage Technology*: The storage technology should store any type of data, and it should be easily scalable so you can add more storage as needed.

- *Governance:* The Assam State Datalake would be governed by Centre for Data Management (CDM). CDM would have a team of Data analysts and Data scientists who would work with various departments in a shared model to provide insights.

- *Data Lineage and data auditing*: Data Lineage deals with data's origins and where it moves over time and what happens to it. This eases errors of corrections in a data analytics process from origin to destination. Data Auditing helps in tracking changes to important data elements, and should capture how/when/and who changes to these elements.

- *Data Purging:* Datalake must have the following considerations for data purging:
  - Data trimming operations like removing the spaces between text data must be done periodically.
  - Check for duplicate files such as images must be run regularly.
  - Data governance and data retention policies must be reviewed regularly.

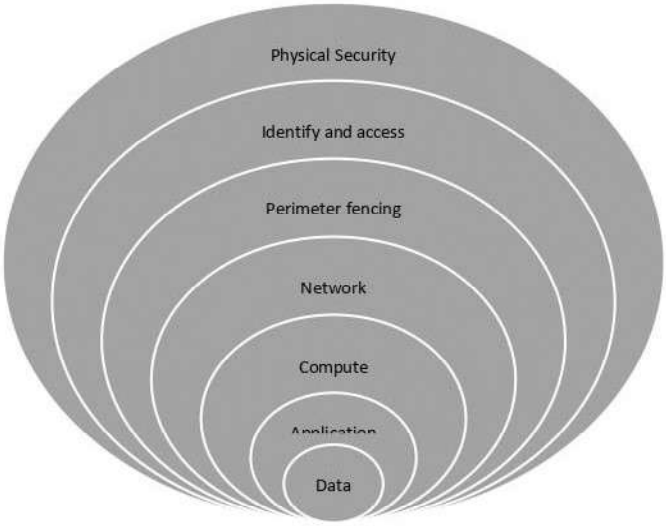
The detailed functional and technical design of the Assam State Datalake will carried out as a downstream activity before the actual implementation is initiated.

The datasets to be archived for long term would be using the Datalake as the central repository. Departments would continue to be the custodians of their respective datasets; and access to these archived datasets would also be governed by the rules of access of permissioned access data. No datasets would be declassified as open data without written request from the concerned department and explicit approval of Seniormost Secretary of the respective Department, as governed by the workflow defined in the above sections.

6.8 **Securing the Data System**:

The state's data would be held across the various information systems by the government departments and autonomous bodies. Therefore the data security guidelines of the information systems would be stricter and would consist of the below cited standards. On the other hand, state data analytics portal would hold mostly the open datasets, hence slightly relaxed data security protocols could be considered here.

Any data system which would be holding any of the three types of data – open access data, permissioned access data or non-shareable (sensitive) data must have various layers of security:



- *Physical Security*: Physical security refers to strict controls over physical access to the areas where government data is stored. Datacentres holding the government data must have extensive layers of protection: access approval, at the facility's perimeter, at the building's perimeter, inside the building, and on the datacentre floor.

- *Identity and access*: This layer of security refers to the IAM controls which are being maintained for any user for accessing the Government data. Use of single sign-on and multi-factor authentication are recommended.

- *Perimeter fencing (Logical):* These are the various security controls being maintained for protecting the data from network-based attacks. This would include the perimeter firewalls being maintained to identify and alert against the malicious attacks. Also, the various other mechanisms to avoid DDoS attacks.

- *Networking:* The use of software-defined networking[9] (SDN) for limiting communication between resources by dividing the networks into micro subnets, virtual networks, implementing policy of zero trust for services trying to access the network resources, etc.

- *Compute:* Malware, unpatched systems, and improperly secured systems open your environment to attacks. The focus in this layer is on making sure the compute resources are secure, and use the proper controls in place to minimize security issues. This could be done by secure access to Virtual machines as well as physical servers, and implement endpoint protection and keep systems patched and current.

- *Application*: The security aspect should be an integral part of application development lifecycle which will help in reducing the vulnerabilities in the code. The application should be audited on a periodic basis to check for OWASP top 10[10] vulnerabilities.

---

[9] Software-defined networking technology is an approach to network management that enables dynamic, programmatically efficient network configuration in order to improve network performance and monitoring, making it more like cloud computing than traditional network management.

[10] OWASP Top 10 is a standard awareness document for developers and web application security. It represents a broad consensus about the most critical security risks to web applications.

- *Data*: This layer of protection is for securing the data which can be ensured by controlling access to data only to the group of people who need it, implemented through access to database.

Security Audit

The data systems collecting, receiving, and archiving data must undergo periodic security audit from a CERT-In empanelled security auditor. All the vulnerabilities identified during the audit must be fixed by the Data Owner/Data Processor department within a reasonable period of time.

Assam Cyber Security Policy

The Data/Information systems would be further guided by the Security guidelines of the Assam Cyber Security Policy, 2020.

6.9 Data Quality and usability assurance

Data owners from the Government of Assam departments, associated institutions / agencies as well as body corporate which routinely collects data in various ways must institutionalise a system of Data Quality Assurance and must have a plan for doing so. Data captured should be devoid of any contaminations, which usually happens when a) process or phenomenon for which data is being recorded, other than one's area of interest, affects the variable value, and b) due to Erroneous values. Below are the two common types of data errors,

- Errors of Commission which happen due to incorrect or inaccurate data entered/recorded. This could occur due to malfunctioning equipment or mistyped data.
- Errors of Omission which happen due to data or metadata not being recorded. This could occur due to inadequate documentation, human error, anomalies in the field.

Data Quality Control/Quality Assurance Strategies (recommended)

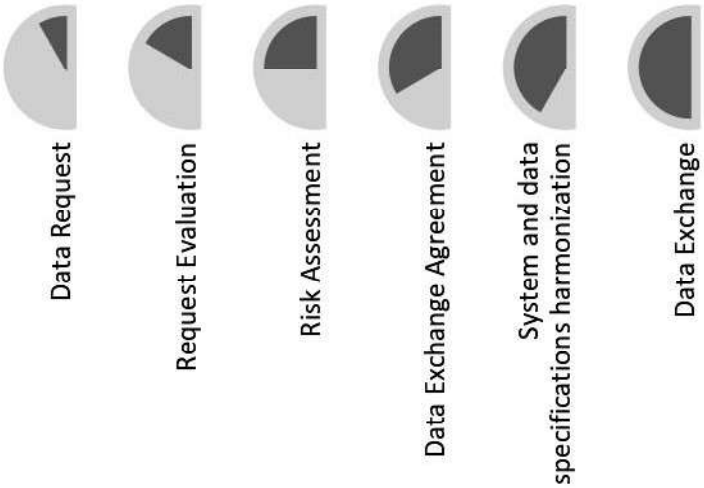
- QA/QC before collection: Define and enforce standards for formats, codes, measurement units, metadata as well as assign responsibility for data quality audit.
- QA/QC during data entry: Implementation of strategies like Double entry which include data keyed in by two independent people, check for agreement with computer verification, record a reading of data and transcribe from the recording, use of text-to-speech program to read data back. However, this may be implemented only if a possibility exists.

  Besides the above, data storage must be designed well so as to minimize the number of times items need to be entered repeatedly, usage of consistent terminology, atomization of data (one cell per piece of information), document changes to data, avoidance of duplicate error checking, allowance of undo if necessary.
- QA/QC after data entry: Some of the checks to be implemented include data line up in columns, no missing, impossible values or anomalies.

  It is also important to look for outliers which are extreme values for a variable using the various statistical models such as normal probability plots, regression, scatter plots, usage of maps, and subtracting values from the mean.

6.10 **Rules of Access to non – open Data**

The Data Exchange provides a component for sharing the data which are designated as non-open. It will be governed by overarching framework consisting of data exchange process, its key components and the overarching governance rules.



Data Exchange Framework

6.10.1 **Data Exchange Standards**

Government of Assam departments, associated institutions / agencies as well as body corporate must exchange and share data in accordance with the requirements set out in this standard. Third parties or external agencies associated with the usage of Government Data will also need to adhere to the requirements of this standard.

Requirements

When exchanging data, the various institutions / actors must follow a minimum set of protocols:

Data Requestor:

- Details to be included in request for data:
  - The purpose and background context for the data request
  - A clear description of the data required.
  - How the data will be used.
  - Whether the data will be shared or distributed and to whom.
  - Whether the request is one-off or on-going and under what conditions the data will be exchanged and managed.

Data Provider:

- To evaluate all the data requests to assess whether the department has the right (or authority to exchange the data requested including:
  - Legislative authority or obligation to share under legislation (Acts) relevant to department or portfolio.
  - If the provider is not the owner of the data, whether there is:
    - A commercial agreement

- Personal individual consent
- Data asset owner's consent (if the data is owned by another department or agency)

- Evaluate all data requests to assess whether the department is ready to exchange the data requested including:
  - Carrying out a risk assessment to determine risk to the department, the Government of Assam and the public
  - Ensuring that where 'sensitive' data is involved, a privacy impact assessment is conducted to ensure that reasonable steps have been taken to protect the data from misuse or loss and unauthorized access, modification, or disclosure.
  - Ensuring data is de-identified wherever possible, unless identified data is essential to enable the data to be fit-for-purpose
  - Assessing whether the provider and requestor have the appropriate processes, technology and infrastructure in place, and sufficient capabilities and capacity to undertake the exchange
  - Whether the data is of sufficient quality to be fit-for-purpose, and if not, to provide appropriate disclaimers as to its use.

- Disclose 'sensitive' information only to the extent required to meet the objectives of the request.

- Ensure that all data exchanges are accompanied by a data exchange arrangement – legally binding, non-legally binding or an Open Data license. The type of arrangement used should be based on who the department is exchanging data with, the level of data protection required, and the level of risk associated with the data and data exchange.

- Ensure all legally binding and non-legally binding data exchange arrangements include these minimum requirements.

- Exchange data to the maximum extent possible under an Open Data license and release State OGD portal as open data unless restricted for reasons of privacy, public safety, security and law enforcement, public health and compliance with the law.

CDM

- Record all requests and data sharing arrangements into a register of data exchange initiatives for government probity and transparency.

- Ensure data exchange arrangements comply with the requirements for managing government and public sector data under these guidelines.

- Ensure all data exchanges are authorized by an officer of the Department at the level commensurate to the risk associated with the data and in accordance with the government's policies and standards notified by various acts and policies.

- Appoint an owner and custodian in each of the requestor and provider organizations who will be accountable and responsible for the data exchange.

Need for Data Exchange Agreement – Formal and Informal

The exchange of permissioned access data must be accompanied by documented data arrangement (formal or informal) which will depend on parameters as outlines in the table below depending on the nature of requests:

Requestor	**Internal**	**External**					
		Within the Government of Assam			Outside the Government of Assam		
	Refers to requestors that are internal to the provider organization	Refers to Government of Assam departments, associated institutions / agencies as well as body corporate which are legal entities in their own right. A legally binding agreement should be entered into with a statutory body when warranted by the associated level of risk.			Refers to all other entities including government funded entities outside of the government, local government, central government, governments in other jurisdictions and any non-government entities.		
Data risk	All levels (1)	Not sensitive (1)		Sensitive	Not sensitive (1)	Sensitive	
Arrangement type	Non-legally binding	Non-legally binding	Legally binding	Non-legally binding	Legally binding		
Format	Informal	Informal	Formal	Formal	Formal	Formal	
Example of such an arrangement	Email	Email	License such as Open Data License (2)	MoU (or other formal non-legally binding mechanism)	License such as Open Data License (2)	Legal Agreement	

.10.2 **Risk Assessment Model**

A risk-based approach to the data request evaluation is recommended, which aims to balance the risk of disclosure with the proposed benefits and outcomes of the initiative being undertaken by the Requestor(s). This assessment should be carried out before the data exchange agreement is being made between the Data Requestor and the Data Provider. The evaluation process involves undertaking a risk assessment using a risk assessment model which evaluates the various risks from the five themes– a) Projects b) People c) Settings d) Data e) Outputs.

Risk assessment Questionnaire

Below is a toolkit which could help us in completion of this evaluation exercise. The queries associated with each aspect will help us identify the risks associated with that theme.

Sl.No.	Theme	Assessment Questions
1.	Project	Is this use of the data appropriate? • Refers to the legal, morals, and ethical considerations surrounding the use of the data. • Are the objectives, outputs, benefits, and outcomes of the initiative reasonable and in alignment with the purpose and functions of the Requestor organization? • What are the risks of loss, harm or detrimental impact to the department, individuals, wider government, general public of sharing (or not sharing) the data? Are there any mitigations?
2.	People	Is the user authorized to access and use the data? • Refers to the knowledge, skills and incentives of the users using the data. • Is the Requestor organization reputable and trustworthy? • Do the staff possess the knowledge (e.g. skills and experience) to effectively use the requested data for the proposed purpose? How will the Provider or Requestor ensure that their staff have appropriate and sufficient knowledge? • What are the roles and responsibilities for all the staff (or user groups) who will have access to the data and what level of access will they have?
3.	Settings	Does the access environment prevent unauthorized use? • Refers to the controls on the way the data is accessed, including physical, procedural and compliance controls. • Does the Requestor possess the technical requirements (e.g. equipment, software), governance policies, and processes to effectively manage and enable the use of the requested data for the proposed purpose? • Where will the data be stored and used? • What security and technical safeguards are in place to ensure data remains secure and protected from unauthorised access and use (e.g. governance, physical safeguards, personnel, and cyber security arrangements)? Safeguards must align with the classification of the data being shared. • How will the data be dealt with after it has been used for this purpose?
4.	Data	Has appropriate and sufficient protection been applied to the data?

		• Refers to whether the data itself contains sufficient information for confidentiality to be breached? • Is 'sensitive' data requested? Is the data required to remain identified? • If not, the Provider should ensure that the data is de-identified. • If identified data is required, the Requestor should outline how they will de-identify the data and ensure that confidentiality is maintained. • If the data is going to be joined or integrated with other datasets, how will this happen and how will the resulting data be used? Does this increase the risk of disclosure? • Are there any potential data quality, matching, reconfiguration, interpretation, or other issues regarding the data being requested?
5.	Outputs	Are the analytical results non-disclosive i.e. Individuals or groups cannot be re-identified from the outputs? This is the final check on the information before it is released which aims to reduce the risk of disclosure to a minimum. • Will the results of the data or analytics work on the shared data be published or disclosed? If so, what is the nature of the proposed publication or disclosure? • Who will be the audience for the publication or disclosure? • What is the likelihood and the extent to which the publication or disclosure may contribute to the unauthorised identification of a person in the data.

Besides the above themes, the risk assessment needs to be completed for other associated risks:

Sl.No.	Theme	Assessment Questions
1.	Reputational risk	Refers to whether there are any: • Threats or danger to the good name or standing of the department. • Risk that the outputs or results of the initiative could contradict or refute any government-wide policy or directive.
2.	Public risk	Refers to whether there are any risks to the safety, security and/or well-being of the general public.

Rating of the risks

The risks associated will be rated from two different perspectives:

a) Likelihood of occurrence
b) Consequence of the risk

Final rating of the risk is usually arrived by referring to the below table:

	Consequence Rating					
Likelihood	**Rating**	**1**	**2**	**3**	**4**	**5**
Almost certain	**5**	Medium	Medium	Significant	High	High
Likely	**4**	Low	Medium	Significant	Significant	High
Neutral	**3**	Low	Medium	Medium	Significant	Significant
Unlikely	**2**	Low	Low	Medium	Medium	Medium
Rare	**1**	Low	Low	Low	Low	Medium

The meaning associated with the consequence rating with respect to various themes and the two other associated risks are cited below:

Consequence Rating				
Insignificant - 1	**Minor – 2**	**Moderate -3**	**Major – 4**	**Catastrophic - 5**
Projects				
No identified ethical aspects or not using data involving people	Having minor ethical risks which can be mitigated, or using highly aggregated or obfuscated data which has no residual personal information	Having ethical risks which require monitoring, or using lightly aggregated or obfuscated data with a possible risk of reidentification or individual information	Having identifiable ethical risk which require significant attention, or using lightly aggregated or obfuscated data with a plausible risk of reidentification of individual information	Clear ethical risks, or using personal information without appropriate deidentification or security controls
People				
Authorized people interacting with the data have the knowledge and skills for required management and use of the data	Authorized people interacting with the data have reasonable knowledge and skills for required management and use of the data	Authorized people interacting with the data have minimal knowledge and skills for required management and use of the data	Authorized people interacting with the data have little or no knowledge and skills for required management and use of the data	Unauthorised management or use of the data
Data				
No sensitive data requiring treatment	Unauthorised disclosure or sensitive data	Unauthorised disclosure of sensitive data	Unauthorised disclosure of sensitive data	Unauthorised disclosure of sensitive data

	to an internal party	to a single external party (not including the general public)	to multiple external parties (not including the general public)	to the general public
Settings				
System accessed with multi-factor user authentication, active action, logging, full audit trail of data lifecycle, anomaly detection, prevention of on-sharing	System accessed with multi-factor user authentication, user action logging, prevention of on-sharing	System accessed with multi-factor user authentication, no ability to readily on-share	System accessed with named user login authentication, limited ability to on-share	System accessed with no restriction on who can access data with ability to on-share
Outputs				
Projects based on open data or projects considered to be highly safe.	Project based on low value data or projects which are considered to be Safe	Projects based on moderate value data or projects which are considered to have a Moderate level of safety	Projects based on high value data or projects which are considered to have a Low level of safety	Projects based on very high value data or projects which are considered not safe.
Reputational Risk				
Minor, adverse local public or media attention or complaints	Media attention of local concern	Significant adverse attention by media and/or public	Serious public or media outcry (State coverage)	Serious public or media outcry (national coverage)
Public Risk				
No public risk to wellbeing or safety or members of the public identified	Minor public risk to wellbeing or safety. Potential for a person to be identified	Significant public risk to well-being or safety. Potential for a person to be identified.	Major public risk to wellbeing or safety or members of the public identified	Serious public risk to wellbeing or safety or members of the public identified.

6.10.3 **Institutional arrangements between Government of Assam agencies**

There will be institutional arrangements between the departments for sharing of permissioned datasets among them. The access to these datasets would be guided by a *Data Exchange Agreement* for usage of these datasets. The components to be included in the agreement would include the following:

Requirement	**Description**	**Arrangement type (Applicability)**		
		Non-legally binding		**Legally binding**
		Informal (e.g, email)	**Formal (e.g., MoU)**	**Formal (Legal Agreement)**
Purpose	The purpose of the initiative that underpins the data exchange, including the associated outputs, benefits, and outcomes achieved and how the data will be used to achieve these benefits and outcome	Yes	Yes	Yes
Background	Context around the initiative and the basis for the data exchange including relevant statutory powers, government policies, operational needs, and organizational strategic directives.	Yes	Yes	Yes
Period of agreement	Commencement date of agreement, how long the	Yes	Yes	Yes

Requirement	Description	Arrangement type (Applicability)		
		Non-legally binding		Legally binding
		Informal (e.g, email)	Formal (e.g., MoU)	Formal (Legal Agreement)
	agreement will be in place and/or the end date			
Key Contacts	Names, roles, and contact details of the appointed representatives of each party of the data exchange arrangement	Yes	Yes	Yes
Obligations	The roles and responsibilities of each party, governance structures in relation to the arrangement and that all appropriate authorizations have been sought. This may include principles around data exchange	Conditional (1)	Yes	Yes
Data description	Description of the data being exchanged, including data types, timeframes (e.g. data from 2010 -2018, broken down by month), data-related standards used (e.g. metadata	Yes	Yes	Yes

Requirement	Description	Arrangement type (Applicability)		
		Non-legally binding		Legally binding
		Informal (e.g, email)	Formal (e.g., MoU)	Formal (Legal Agreement)
	standards, GIS standards, industry standards), data security classification, whether the data has been deidentified, and the method used)			
Terms of use and disclosure	How the data will be used, joined, or integrated, de-identified for privacy, reproduced, published internally, externally, or not at all, or commercialized. With whom may be requestor share or distribute the data or outputs resulting from using the data and under what conditions this may occur	Conditional (1)	Yes	Yes
Intellectual Property (IP) and licensing	Who has ownership of the data and IP rights? Who will own any new IP developed? Can the requestor use the data for commercial	No	Yes	Yes

Requirement	Description	Arrangement type (Applicability)		
		Non-legally binding		Legally binding
		Informal (e.g, email)	Formal (e.g., MoU)	Formal (Legal Agreement)
	purposes or building a brand or reputation and under what conditions? This may include an Open Data License and associated attribution			
Data quality statement	Details of the quality of data. A data quality statement will be provided in the first instance and updated when there are changes to any data quality dimensions, in accordance with the best practices.	Conditional (1)	Yes	Yes
Data exchange and management	How the data will be transmitted (methods and standards) by the Provider to the Requestor, how the data will be managed by the Requestor, including data security and privacy (including de-identification)	Yes	Yes	Yes
Service levels	Service levels around the	Conditional (2)	Conditional (2)	Conditional (2)

Requirement	Description	Arrangement type (Applicability)		
		Non-legally binding		Legally binding
		Informal (e.g, email)	Formal (e.g., MoU)	Formal (Legal Agreement)
	provision of data including data availability and reliability targets, maintaining data quality (including de-identification of data), complaints handling process and response times and consequences of not meeting service level targets.			
Change management	The process for managing changes to the data provided – what, how, who and when this is communicated from the Provider to the Requestor	No	Conditional (2)	Conditional (2)
Data retention/ disposal	Relevant data retention periods and whether the data should be returned to the Provider or disposed of at the end of the retention period	No	Yes	Yes
Breach in data use or disclosure	Outline the process for managing	Conditional (1)	Yes	Yes

Requirement	Description	Arrangement type (Applicability)		
		Non-legally binding		Legally binding
		Informal (e.g, email)	Formal (e.g., MoU)	Formal (Legal Agreement)
	unauthorized use or disclosure of data and any sanctions for failure to comply.			
Fees or charges	Outline any fees or charges that apply for providing the data and payment terms and conditions. This will apply in cases which has been approved by CDO.	No	No	Yes
Compliance	The Provider's rights to monitor compliance with the exchange standards and terms of agreement.	No	No	Yes
Review of arrangement	If the arrangement is a rolling arrangement, there should be a date to review its ongoing effectiveness	No	Conditional (2)	Conditional (2)
Schedules to the arrangement	A separate schedule should be provided for each dataset exchanged and should include the dataset name,	No	Conditional (3)	Conditional (3)

Requirement	Description	Arrangement type (Applicability)		
		Non-legally binding		Legally binding
		Informal (e.g, email)	Formal (e.g., MoU)	Formal (Legal Agreement)
	description, data owner (or custodian), data fields, data definitions, data quality statement and data security classification.			

Legends:
Conditional (1) applies if personal or sensitive information or confidential data are involved
Conditional (2) applies if data provision is recurring (i.e. not one-off)
Conditional (3) applies if more than one dataset is exchanged.

6.10.4 **Implementation of Data sharing using the Data Exchange (DE) tool**

Once the Data request has been made by the Data Requestor and the various data sharing agreements has been made after rigorous risk assessment by the Data Provider, the identified datasets are shared through the State Data Exchange (DE). The below are the guidelines which define the various processes associated with sharing the data across the various government/non-government actors based on the type of dataset. The requestor and provider would be making necessary arrangements in the consumer app and the provider app respectively, for enabling the sharing of the data.

- Dataset On-boarding: The open access or permissioned access datasets could be on-boarded into the DE for sharing across the Government of Assam departments, associated institutions / agencies as well as body corporate after necessary pre-processing.
  - Data Quality assessment must be done to check for completeness and consistency of the dataset. Dataset must have followed the metadata standard as set by CDM.
  - Datasets shared must be consumable such that data sets as shared by data processor are available in an electronically sharable format and the ease by which data of variable size and formats is allowed to be consumed.

  Data on-boarding after necessary data validation will include the below steps:
  - Segregation of Dataset into Static and Dynamic data, which helps in removing redundancy that can occur during the data ingestion and also reduces the growth in size of the data. The data components which seldom change and need not be ingested often are marked as static, and the rest

of the data components are marked as dynamic. The API structure is changed according to this segregation.

- Ingestion ready data is then catalogued, with the help of metadata which comes with the dataset. And the dataset is now ready to the consumed by the external apps from the various information systems.

- Access to Dataset: The sharing of datasets across the Government of Assam departments, associated institutions / agencies as well as body corporate would be governed by general guidelines set by CDM for open-access data, permissioned access data and sensitive data. However, the consumer application of the Data Requestor would follow the below protocols to access the requested dataset from the provider application of the Data processor.
  - Discover Data: a consumer application uses the search APIs of the catalogue service of DE to find interested entities. Once the entities are identified, a consumer App can request for consent and access their data. In case, the data exchange agreement is already in place, the request for consent may not be needed.
  - Request authorization and Access Data: A consumer application can request access to dataset from an DE compliant resource server using any one of the supported APIs. If requested dataset does not require authorization that is, does not have an authorization policy, or the request contains a valid access token, then the resource server serves the request after token validation.

    If data requires authorization and the request does not contain a token, the resource server initiates authorization following the API access protocol. The protocol must support the UMA 2.0[11] workflows:
    a) Accept policies specified in a policy language.
    b) Support identities based on access certificates issued by DE certifying authority (CA)
    c) Accept claims based on access certificates issued by a set of trusted CAs
    d) No support interactive claims gathering. All claims shall be pushed.
    e) Include a reference to the policy object in access tokens issued
  - Revoke of access: After the consumer app has access to the data, the provider shall be able to revoke access to a particular resource by calling the revoke API.

6.10.5 Enabling Data sharing by Consumer and Provider App

The requestor and the provider organizations must make proper assessment in their respective Information Systems (Data Systems) for enabling the data sharing. The below aspects must be assessed by the respective technical teams before initiating the activity of data sharing.

Performance Impact Assessment

[11] UMA 2.0: User- managed Access (UMA) 2.0 is a federated authorization standard protocol built on top of Open Authentication (OAuth) 2.0 enables party-to-party sharing.

Both the organizations must make the performance impact assessment of the source and the target system for extraction, transmission and receipt of the dataset. Also, the technical teams must ensure that necessary performance testing is carried out before the initiation of the data sharing process. Also, a plan must be devised for mitigating the impact (e.g. transfer only outside the working hours).

Storage of Data

The storage capacity assessment must be carried out to understand the impact especially when the data exchange is an ongoing activity, and involves large data files. It must also be ascertained if there is a need for data encryption when it is stored. In case, the data is requested only for a short period of time, it must also be assessed as to how the data will be disposed of.

Data Security

It must be ascertained if there are adequate security controls to ensure that it is protected from unauthorized access, modification and loss. The access controls associated with read and write permissions among the users (or groups) may be assessed. Some modalities must be set at both the consumer and provider apps for dealing with interruptions or faults within the data exchange. These should include the monitoring of errors and faults, and ability to restart while ensuring no data is lost and no duplicates are produced.

6.10.6 **Operationalizing the Data exchange**

Operationalization of State Data Exchange occurs whenever there is an exchange of actual (real, not test) data. The Data exchange occurs between the Data Requestor (via the Consumer app) and the Data Provider (via the Provider app). This may occur as a one-off or recurring process. Below table has the key considerations when operationalizing a data exchange which are applicable both as one-off as well as recurring data exchanges:

Sl.No.	Area	Considerations	Provider	Requestor
1	Change management (only applicable to recurring data exchanges)	Changes may occur from time to time that impact data quality (such as the way data is collected, how dimensions are defined, how measures are calculated, data stops being collected) or impact the format or method of the exchange. • If the changes impact data quality, the Provider should update and reissue the data quality statement to inform the data requestor of the changes. • If the changes impact metadata, the provider should update the metadata and	Yes	N/A

Sl.No.	Area	Considerations	Provider	Requestor
		inform the Requestor of the update. • If changes impact the format or method of the exchange, the Provider should work with the Requestor to test the new format or method. • Change management should be addressed as part of the data agreement arrangement. • Notification of changes should be within the timeframe set out in the arrangement or at least when the next tranche of data is provided.		
2	Exception handling	Exception handling is required when there is a failure in the data exchange process. This can happen at any point of the process, from data collation to transmission. An example is when system outages (such as source systems of the data, data distribution portals) occur that delay the provision of the data: The Provider should consider how the failure is going to be remediated. Exception handling (including complaints handling) should be addressed as part of the arrangement in the service levels section.	Yes	N/A
3	Contract management (including service levels)	As a part of the arrangement, both parties should appoint a data exchange owner and custodian. • The owner will be accountable and have the power to authorize or stop the exchange, if required. • The custodian will be the contract manager and main operational point of contact for the exchange.	Yes	Yes

Sl.No.	Area	Considerations	Provider	Requestor
		The custodian should monitor each party's obligations as set out in the arrangement and manage issues if they arise. This may include: • Monitoring and reporting performance against the obligations (such as service levels) • Ensuring changes and exceptions (see above) are managed. • Liaising with the other party to resolve any issues. • Monitoring changes to the terms and conditions of the arrangement and amending the arrangement where necessary. • Renegotiating or extending the contract, if required. Complex, high risk data exchanges may need a regular face to face status meeting.		
4	Monitoring and reporting	As part of contract management, custodians will need to be able to monitor each party's obligations of the arrangement. This may require reviewing or updating various reports such as: • Monitoring log reports of when data is provided to ensure data is provided within the timeframe set out in the arrangement. • Monitoring log reports of when data is accessed and by whom to ensure data is only accessed by permitted users • Monitoring system error log reports to identifying system failures • Updating the data exchange register for the data exchange(s)	N/A N/A Yes Yes Yes Yes	Yes Yes Yes Yes Yes Yes

Sl.No.	Area	Considerations	Provider	Requestor
		• Monitoring data exchange risks and updating the risk register • Monitoring the sharing, management, and security of the data in accordance with these guidelines. • Where applicable, reviewing the output to ensure deidentification has occurred	Yes	Yes
5	Resourcing	Do all the people involved in the data exchange understand their roles and responsibilities? • The roles that may be involved in an exchange are listed below. The roles do not always have to be held by different people. An individual may sometimes have multiple roles: o Data exchange owner o Data exchange custodian o Information management specialists o Information technology specialists (security, ETL, data warehousing, testing) o Data analysts (analysts, researchers) o Data consumers (internal business managers, general staff, external: other public sector, private sector, special interest groups, general public) • Do all the people involved in the data exchange have sufficient knowledge and skills to manage and or use the data for its intended purpose?	Yes Yes	Yes Yes

Sl.No.	Area	Considerations	Provider	Requestor
		If not, education and training may be required to be provided.		

6.10.7 **Data Owners for the Departmental Registries**

Single sources of Truth: Several Government of Assam departments, associated institutions / agencies as well as body corporate may have overlapping and contradictory and/or outdated data (master data), as far as possible, a Government Department(s)/ Organization(s) seeking data under this Policy should seek the same from such Government Department(s)/ Organization(s), which in its reasonable opinion is the owner of that data as sought by them.

The Centre for Data Management (CDM) would in due course of time, notify the Master(s) of Data which would provide, inter alia, the Data Owners and the nature and fields of data for which a Department/ Data Owners shall be considered to be a single source of truth, when such Master(s) of Data are notified; and such entities must maintain this master data as registries. Each department must ensure that it seeks the required data, for which Master Data exist, only from such Master(s) of Data.

Below are the steps to be taken for notifying a Data Storage as a single source of truth:

Steps to be taken	Creator	Checker	Approver
Step 1: The Departmental Nodal officers creates a request for establishing the departmental masters storing data fields specific to that particular Government of Assam department, associated institution / agency as well as body corporate, as single source of truth of Beneficiary information for all departments.	Departmental Data Officer	CDO	ASDM Steering Committee
Step 2: Notification from the department on designating registry as a single source of truth.	Departmental Assistants	Departmental Data Officer	Departmental Data Custodian/ Seniormost Secretary of the Department.

7 Glossary of Terms

Sl. No.	Term	Details
1.	Data	facts and statistics collected together for reference or analysis
2.	System	a set of things working together as parts of a mechanism or an interconnecting network; a complex whole
3.	Digital	data expressed as series of the digits 0 and 1, typically represented by values of a physical quantity such as voltage or magnetic polarization.
4.	Digitize	convert (pictures, text, or sound) into a digital form that can be processed by a computer
5.	Transform	make a marked change in the form, nature, or appearance of.
6.	Data Systems	Data system is a term used to refer to an organized collection of symbols and processes that may be used to operate on such symbols. Data Systems, in the context of this document, also means computer, electronic or telecommunications or network systems of any variety (including data bases, websites, hardware, software, storage, switching and interconnection devices and mechanisms, whether on-premises or provided as a service by a third party). These are also known as the Information Systems.
7.	Digital Transformation	Digital transformation is the process of using digital technologies to create new — or modify existing — service delivery processes, culture, and citizen experiences to meet the obligations of the government towards the citizens.
8.	Interoperability	the ability of data systems (computer systems or software) to exchange and make use of information
9.	Data Democracy	Data democratization is when an organization makes data accessible to all employees and stakeholders, and educates them on how to work with data, regardless of their technical background.
10.	Data Management	Data management is the effective practice of collecting, storing, protecting, delivering, and processing data.
11.	Data Governance	Data governance refers to the collection of practices, policies, and roles related to the effective acquisition, management, and utilization of data—ensuring that the data provides as much value as it can within an organization.

Sl. No.	Term	Details
12.	Data Analytics	Data Analytics is used for the discovery, interpretation, and communication of meaningful patterns in data. It also entails applying data patterns toward effective decision-making.
13.	Metadata	a set of data that describes and gives information about other data.
14.	Machine-readable format	Machine-readable data, or computer-readable data, is data in a format that can be processed by a computer. Machine-readable data is preferably the structured data.
15.	Dataset	A data set (or dataset) is a collection of data. In the case of tabular data, a data set corresponds to one or more database tables, where every column of a table represents a particular variable, and each row corresponds to a given record of the data set in question.
16.	Data infrastructure	A data infrastructure is a collection of data assets, the bodies that maintain them and guides that explain how to use the collected data.
17.	Data derivatives	data derivative is knowledge or information that is inferred or derived from a data set, based on patterns mined by means of computational techniques such as clustering, association rules, regression analyses, neural networks, reinforcement learning, unsupervised algorithms and the more.
18.	Data privacy	Data privacy is the right of a citizen to have control over how their personal information is collected and used.
19.	Intellectual property rights	Intellectual property rights are the rights given to persons over the creations of their minds.
20.	Data access	Data access is a generic term referring to a process which has both an IT-specific meaning and other connotations involving access rights in a broader legal and/or political sense. In the former it typically refers to software and activities related to storing, retrieving, or acting on data housed in a database or other repository. The later refers to data being available to be access by any person or organization based on the IPR associated with the data.
21.	Data Archive	The digital location where machine-readable data is stored, worked upon / analysed, documented prior to a cut-off past date.

Sl. No.	Term	Details
22.	Data Generation	Initial collection of data or subsequent addition of data to the same specification. This may be data specifically collected for a particular objective or may be a consequence of the authorised administrative processes of the Government.
23.	Creative commons license	A Creative Commons (CC) license is one of several public copyright licenses that enable the free distribution of an otherwise copyrighted "work". A CC license is used when an author wants to give other people the right to share, use, and build upon a work that the author has created. CC provides an author flexibility (for example, they might choose to allow only non-commercial uses of a given work) and protects the people who use or redistribute an author's work from concerns of copyright infringement as long as they abide by the conditions that are specified in the license by which the author distributes the work.
24.	Application Programming Interface (API)	An application programming interface (API) is a way for two or more computer programs to communicate with each other. It is a type of software interface, offering a service to other pieces of software. A document or standard that describes how to build or use such a connection or interface is called an API specification.
25.	Internet to Things (IOT)	The Internet of things describes physical objects with sensors, processing ability, software and other technologies that connect and exchange data with other devices and systems over the Internet or other communications networks.
26.	Geo-spatial data	Geospatial data (also known as “spatial data”) is used to describe data that represents features or objects on the Earth's surface.
27.	Big Data	Big data is the one associated with large body of information that we could not comprehend when used only in smaller amounts. In it primary definition though, Big data refers to data sets that are too large or complex to be dealt with by traditional data-processing application software. Big data is a combination of structured, semi-structured and unstructured data collected by organizations that can be mined for information and used in machine learning projects, predictive modelling and other advanced analytics applications.

Sl. No.	Term	Details
28.	Personal Health Information	Personal health information (PHI) is a category of information that refers to an individual's medical records and history, which are protected under Acts such as the Health Insurance Portability and Accountability Act (HIPAA).
29.	Bots	An Internet bot, web robot, robot or simply bot, is a software application that runs automated tasks over the Internet, usually with the intent to imitate human activity on the Internet, such as messaging, on a large scale.
30.	Drones	A drone is an unmanned aircraft. Drones are more formally known as unmanned aerial vehicles (UAVs) or unmanned aircraft systems. Essentially, a drone is a flying robot that can be remotely controlled or fly autonomously using software-controlled flight plans in its embedded systems, that work in conjunction with onboard sensors and a global positioning system (GPS).
31.	Aerial Imaging	Aerial photography (or airborne imagery) is the taking of photographs from an aircraft or other airborne platforms. When taking motion pictures, it is also known as aerial videography. Platforms for aerial photography include fixed-wing aircraft, helicopters, unmanned aerial vehicles (UAVs or "drones"), balloons, blimps and dirigibles, rockets, pigeons, kites, or using action cameras while skydiving or wing suiting. Handheld cameras may be manually operated by the photographer, while mounted cameras are usually remotely operated or triggered automatically.
32.	Master Data	Master data represents "data about the business entities that provide context for business transactions".
33.	Transaction data	Transaction data, or transaction information, constitute a category of data describing transactions. Transaction data/information gather variables generally referring to reference data or master data – e.g. dates, times, time zones, currencies. Typical transactions are: o Financial transactions about orders, invoices, payments; o Work transactions about plans, activity records; o Logistic transactions about deliveries, storage records, travel records, etc.

Sl. No.	Term	Details
34.	Management Information Systems (or Information Systems)	A management information system is an information system used for decision-making, and for the coordination, control, analysis, and visualization of information in an organization.
35.	Database	In computing, a database is an organized collection of data stored and accessed electronically. Small databases can be stored on a file system, while large databases are hosted on computer clusters or cloud storage.
36.	Application	An application program is a computer program designed to carry out a specific task other than one relating to the operation of the computer itself, typically to be used by end-users.
37.	Personally Identifiable Information	Data about or relating to a Data Principal who is directly or indirectly identifiable, whether online or offline, or any combination of such features with any other information; and shall include any inference drawn from such data for the purpose of profiling.
38.	Data Warehouse	A Data Warehouse (DW) is a relational database that is designed for query and analysis rather than transaction processing.
39.	Data Lake	A data lake is a centralized repository designed to store, process, and secure large amounts of structured, semi-structured, and unstructured data.
40.	Body Corporate	Body corporate broadly means a corporate entity which has a legal existence. The term "body corporate" is defined in Section 2(11) of the Companies Act, 2013. This includes a private company, public company, one personal company, small company, Limited Liability Partnerships, foreign company etc.
41.	Development schemes	Various programs of social development run by the Union or State Government for fulfilling its obligations as a Welfare State.
42.	Process Improvement	A continual improvement process, also often called a continuous improvement process, is an ongoing effort to improve products, services, or processes. These efforts can seek "incremental" improvement over time or "breakthrough" improvement all at once.
43.	Cross-functional Analysis	Cross-functional analysis refers to the analysis of data across the various management functions to achieve the desired objectives.

Sl. No.	Term	Details
44.	Legacy Data	Old information that an organization has, especially information stored in an old-fashioned way.
45.	Industrial sensors	Sensor/Detectors/Transducers are electronic or electrical devices. These special electronic sensitive materials sense, measure, and detect changes in the position, temperature, displacement, electrical current, and multiple parameters of industrial equipment.
46.	Data Exchange	Data exchange is the process of taking data structured under a source schema and transforming it into a target schema, so that the target data is an accurate representation of the source data. Data exchange allows data to be shared between different computer programs.
47.	Data Organization	An organization holding substantial amount of data.
48.	Data Dictionary	A data dictionary contains metadata i.e data about the database.
49.	International Standards Organization (ISO)	The International Organization for Standardization is an international standard development organization composed of representatives from the national standards organizations of member countries. Membership requirements are given in Article 3 of the ISO Statutes.
50.	Data Standards	Data standards are the rules for structuring information collected by the ID system which facilitate semantic interoperability. A set of agreed-upon data standards ensures that the data entered into a system can be reliably read, sorted, indexed, retrieved, and communicated between systems.
51.	ASCII	ASCII, abbreviated from American Standard Code for Information Interchange, is a character encoding standard for electronic communication. ASCII codes represent text in computers, telecommunications equipment, and other devices.
52.	Unicode	Unicode, formally The Unicode Standard, is an information technology standard for the consistent encoding, representation, and handling of text expressed in most of the world's writing systems.
53.	Hexadecimal	In mathematics and computing, the hexadecimal numeral system is a positional numeral system that represents numbers using a radix (base) of 16.
54.	Base64	Base64 is a group of binary-to-text encoding schemes that represent binary data (more specifically, a

Sl. No.	Term	Details
		sequence of 8-bit bytes) in sequences of 24 bits that can be represented by four 6-bit Base64 digits.
55.	MIME	Multipurpose Internet Mail Extensions is an Internet standard that extends the format of email messages to support text in character sets other than ASCII, as well as attachments of audio, video, images, and application programs.
56.	Data compression	Data compression is a reduction in the number of bits needed to represent data. Compressing data can save storage capacity, speed up file transfer and decrease costs for storage hardware and network bandwidth.
57.	gzip	gzip is a file format and a software application used for file compression and decompression
58.	Comma-delimited format	A comma-separated values file is a delimited text file that uses a comma to separate values. Each line of the file is a data record. Each record consists of one or more fields, separated by commas. The use of the comma as a field separator is the source of the name for this file format.
59.	QR codes	A quick response (QR) code is a type of barcode that stores information and can be read by a digital device, such as a cell phone
60.	Electronic Registry	All digital platforms require master data and actor (person/entity/thing) data related to that system be maintained for identification, validation, etc. For example, a property tax system needs to maintain master data about properties, land boundaries, tax codes, tax payers, inspection officers, etc. in a structured and validated fashion so as to help manage the property tax transaction in a seamless manner
61.	Digital ID	Digital Identity allows person to prove who he is online, and he can re-use your Digital Identity whenever you need it.
62.	Federated Access	Federated authentication is also a technology that allows users to access multiple tools, apps, and domains with only one set of credentials. Once the user gets into a system using his log-in credentials, he/she is authenticated by the system and can access multiple resources in the organization without using any other credentials.
63.	Federated Registry	User registry federation is used when user and group information is spread across multiple registries. For example, the information might be in two different

Sl. No.	Term	Details
		LDAPs, in two subtrees of the same LDAP, in a file, or the users are of a system.
64.	Single Sign-On	Single sign-on (SSO) is an authentication scheme that allows a user to log in with a single ID to any of several related, yet independent, software systems.
65.	Credentialing	Credentialing is the process of granting a designation, such as a certificate or license, by assessing an individual's knowledge, skill, or performance level.
66.	Instant messaging platforms	An online channel which allows asynchronous communication between two different individuals using the Internet.
67.	Blockchain	A blockchain is a decentralized, distributed and public digital ledger that is used to record transactions across many computers so that the record cannot be altered retroactively without the alteration of all subsequent blocks and the consensus of the network.
68.	ICT	Information and Communications Technology (ICT) is the convergence of computing, telecommunication and governance policies for how information should be accessed, secured, processed, transmitted and stored
69.	Cloud	Cloud computing is the on-demand availability of computer system resources, especially data storage and computing power, without direct active management by the user. Large clouds often have functions distributed over multiple locations, each of which is a data center.
70.	Virtualization	Virtualization is technology that lets you create useful IT services using resources that are traditionally bound to hardware. It allows you to use a physical machine's full capacity by distributing its capabilities among many users or environments.
71.	Data Center (DC)(SDC/NDC)	A data center or data centre is a building, a dedicated space within a building, or a group of buildings used to house computer systems and associated components, such as telecommunications and storage systems. SDC is the State Data Center maintained by DITECH, and NDC is the National Data Center maintained by NIC.
72.	Private Cloud	A private cloud is a cloud service that is exclusively offered to one organization.
73.	Public Cloud	A public cloud is an IT model where public cloud service providers make computing services—including compute and storage, develop-and-deploy

Sl. No.	Term	Details
		environments, and applications—available on-demand to organizations and individuals over the public internet.
74.	Hybrid Cloud	Hybrid Cloud refers to a cloud computing model that uses a combination of at least one private cloud and at least one public cloud, which works together to provide a flexible mix of cloud computing services. Hybrid cloud computing extends infrastructure and operations consistently to provide a single operating model that manages application workloads across both environments, allowing for seamless migration of workloads from private to or from public cloud as business needs dictate.
75.	Digital Signature	A digital signature is a mathematical scheme for verifying the authenticity of digital messages or documents. A valid digital signature, where the prerequisites are satisfied, gives a recipient very high confidence that the message was created by a known sender, and that the message was not altered in transit. A digital signature is needed for verifying the identity of the issuer as per the IT Act 2000.
76.	Cyber Security	Cyber security is the practice of defending computers, servers, mobile devices, electronic systems, networks, and data from malicious attacks.
77.	Information Security Management System	Information security management defines and manages controls that an organization needs to implement to ensure that it is sensibly protecting the confidentiality, availability, and integrity of IT assets (hardware and software) from threats and vulnerabilities.
78.	Asynchronous Communication	In telecommunications, asynchronous communication is transmission of data, generally without the use of an external clock signal, where data can be transmitted intermittently rather than in a steady stream. Any timing required to recover data from the communication symbols is encoded within the symbols.
79.	Differential privacy	Differential privacy is a system for publicly sharing information about a dataset by describing the patterns of groups within the dataset while withholding information about individuals in the dataset.

Sl. No.	Term	Details
80.	I/O Device	Input and Output. This term is used for representing any device for Input and output of data eg. Keyboard, Monitor, mouse, etc.
81.	Attack surface area	The attack surface of a software environment is the sum of the different points where an unauthorized user can try to enter data to or extract data from an environment. Keeping the attack surface as small as possible is a basic security measure.
82.	Secure defaults	Secure by Default is about taking a holistic approach to solving security problems at root cause rather than treating the symptoms; acting at scale to reduce the overall harm to a particular system or type of component.
83.	Principle of least privileges	In information security, computer science, and other fields, the principle of least privilege (PoLP), also known as the principle of minimal privilege (PoMP) or the principle of least authority (PoLA), requires that in a particular abstraction layer of a computing environment, every module (such as a process, a user, or a program, depending on the subject) must be able to access only the information and resources that are necessary for its legitimate purpose.
84.	Principle of defense-in-depth	Defense in depth is a concept used in information security in which multiple layers of security controls (defense) are placed throughout an information technology (IT) system. Its intent is to provide redundancy in the event a security control fails or a vulnerability is exploited that can cover aspects of personnel, procedural, technical and physical security for the duration of the system's life cycle.
85.	Principle of separation of duties	Separation of duties, also known as segregation of duties is the concept of having more than one person required to complete a task. It is an administrative control used by organisations to prevent fraud, sabotage, theft, misuse of information, and other security compromises.
86.	Security-by-obscurity	Security through obscurity (STO) is a process of implementing security within a system by enforcing secrecy and confidentiality of the system's internal design architecture. Security through obscurity aims to secure a system by deliberately hiding or concealing its security flaws.
87.	Non-repudiation	Non-repudiation refers to a situation where a statement's author cannot successfully dispute its

Sl. No.	Term	Details
		authorship or the validity of an associated contract. The term is often seen in a legal setting when the authenticity of a signature is being challenged. In such an instance, the authenticity is being "repudiated"
88.	Data Provenance	Data provenance is the documentation of where a piece of data comes from and the processes and methodology by which it was produced. It is also one of the principles of IT Act 2000.
89.	Data Archiving	Data archiving is the process of moving data that is no longer actively used to a separate storage device for long-term retention.
90.	Data Purge	Data purging is the process of deleting data from a database. When you purge data, it is permanently erased and cannot be restored.
91.	Authentication Service Agencies (ASA)	ASAs are agencies that have established secured leased line connectivity with the CIDR compliant with UIDAI's standards and specifications. ASAs offer their UIDAI-compliant network connectivity as a service to requesting entities (such as AUAs/KUAs) and transmit their authentication requests to CIDR.
92.	AuA/KuA	Authentication User Agency (AUA) is an entity engaged in providing Aadhaar Enabled Services to Aadhaar number Holder, using the authentication as facilitated by the Authentication Service Agency (ASA). An AUA may be government / public / private legal agency registered in India, that uses Aadhaar authentication services of UIDAI and sends authentication requests to enable its services / business functions. A requesting entity (such as AUA, KUA) connects to the CIDR through an ASA (either by becoming ASA on its own or by contracting services of an existing ASA).
93.	IAM	Identity management, also known as identity and access management, is a framework of policies and technologies to ensure that the right users have the appropriate access to technology resources. IdM systems fall under the overarching umbrellas of IT security and data management.
94.	Single sign-on	Single sign-on (SSO) is an authentication scheme that allows a user to log in with a single ID to any of several related, yet independent, software systems.
95.	Multi-factor authentication	Multi-factor authentication (MFA; encompassing two-factor authentication, or 2FA, along with similar terms)

Sl. No.	Term	Details
		is an electronic authentication method in which a user is granted access to a website or application only after successfully presenting two or more pieces of evidence (or factors) to an authentication mechanism: knowledge (something only the user knows), possession (something only the user has), and inherence (something only the user is).
96.	Security Controls	Security controls are safeguards or countermeasures to avoid, detect, counteract, or minimize security risks to physical property, information, computer systems, or other assets. In the field of information security, such controls protect the confidentiality, integrity and availability of information.
97.	Network-based attacks	Network-based attacks are attacks designed to compromise network security by either eavesdropping on or intercepting and manipulating network traffic. These may be active attacks, wherein the hacker manipulates network activity in real-time; or passive attacks, wherein the attacker sees network activity but does not attempt to modify it.
98.	Firewall	A firewall is a network security device or software that monitors traffic to or from your network. It allows or blocks traffic based on a defined set of security rules.
99.	DDoS	DDoS Attack means "Distributed Denial-of-Service (DDoS) Attack" and it is a cybercrime in which the attacker floods a server with internet traffic to prevent users from accessing connected online services and sites.
100.	Software defined networking	Software-defined networking technology is an approach to network management that enables dynamic, programmatically efficient network configuration in order to improve network performance and monitoring, making it more like cloud computing than traditional network management.
101.	Subnets	A subnetwork or subnet is a logical subdivision of an IP network. The practice of dividing a network into two or more networks is called subnetting. Computers that belong to the same subnet are addressed with an identical most-significant bit-group in their IP addresses.
102.	Zero-trust policy	Zero Trust is a security framework requiring all users, whether in or outside the organization's network, to be authenticated, authorized, and continuously validated for security configuration and posture before being

Sl. No.	Term	Details
		granted or keeping access to applications and data. Zero Trust assumes that there is no traditional network edge; networks can be local, in the cloud, or a combination or hybrid with resources anywhere as well as workers in any location.
103.	Malware	Malware is any software intentionally designed to cause disruption to a computer, server, client, or computer network, leak private information, gain unauthorized access to information or systems, deprive access to information, or which unknowingly interferes with the user's computer security and privacy.
104.	Virus	A computer virus is a type of computer program that, when executed, replicates itself by modifying other computer programs and inserting its own code. If this replication succeeds, the affected areas are then said to be "infected" with a computer virus, a metaphor derived from biological viruses.
105.	Unpatched systems	Unpatched systems or software is a computer code containing known security weaknesses. Unpatched vulnerabilities refer to weaknesses that allow attackers to leverage a known security bug that has not been patched by running malicious code. Software vendors write additions to the codes, known as "patches," when they come to know about these application vulnerabilities to secure these weaknesses.
106.	OWASP	The Open Web Application Security Project is an online community that produces freely-available articles, methodologies, documentation, tools, and technologies in the field of web application security. The OWASP provides free and open resources.
107.	Security Audit	A security audit or information security audit is an audit on the level of information security in an organization. It is an independent review and examination of system records, activities and related documents.
108.	CERT	The Indian Computer Emergency Response Team (CERT-IN or ICERT) is an office within the Ministry of Electronics and Information Technology of the Government of India. It is the nodal agency to deal with cyber security threats like hacking and phishing. It strengthens security-related defence of the Indian Internet domain.
109.	Atomization of data	It refers to the storage of self-contained data units (documents, file etc) on a massively distributed

Sl. No.	Term	Details
		network, and in such a manner that each data unit is split onto a large number of tiny segments (or atoms), the same being atoms which are replicated across the network for later retrieval.
110.	Normal probability plots	The normal probability plot is a graphical technique to identify substantive departures from normality. This includes identifying outliers, skewness, kurtosis, a need for transformations, and mixtures. Normal probability plots are made of raw data, residuals from model fits, and estimated parameters.
111.	Regression	In statistical modelling, regression analysis is a set of statistical processes for estimating the relationships between a dependent variable and one or more independent variables.
112.	Scatter plots	A scatter plot is a type of plot or mathematical diagram using Cartesian coordinates to display values for typically two variables for a set of data.
113.	UMA	User-Managed Access (UMA) 2.0 is a lightweight access control protocol that defines a centralized workflow to allow an entity (user or corporation) to manage access to their resources. UMA 2.0 extends the OAuth 2.0 protocol and gives resource owners granular management of their protected resources by creating authorization policies on a centralized authorization server, such as AM. The authorization server grants delegated consent to a requesting party on behalf of the resource owner to authorize who and what can get access to their data and for how long.
114.	Certifying authority	Certifying Authorities (CA) are the agencies who have been granted a license to issue a digital signature certificate under Section 24 of the Indian IT-Act 2000.
115.	App	A software application, often a small, specialized program often for mobile devices.
116.	Knowledge economy	The knowledge economy is an economic system in which the production of goods and services is based principally on knowledge-intensive activities that contribute to advancement in technical and scientific innovation.
117.	Digital Empowerment of Citizens	Digital India is a programme to transform India into digital empowered society and knowledge economy. This program aims to achieve Digital empowerment of citizens through,

Sl. No.	Term	Details
		The creation of digital infrastructure, Delivery of services digitally and promoting Digital literacy.
118.	Data Governance Quality Index	The Data Governance Quality Index (DGQI) toolkit provides a unique framework for self-assessment of data preparedness levels across the Government of India. DGQI is based on internationally accepted data preparedness assessment models from private and public sectors but appropriately contextualized for India.
119.	Digital Economy	The digital economy is a portmanteau of digital computing and economy, and is an umbrella term that describes how traditional brick-and-mortar economic activities are being transformed by Internet, World Wide Web, and blockchain technologies.
120.	Data Strategy	Data strategy refers to the tools, processes, and rules that define how to manage, analyze, and act upon business data.
121.	Multilingual interfaces	Multilingual User Interface (MUI) enables the localization of user interfaces for globalized applications.
122.	GIGW	Guidelines for Indian Government Websites (GIGW)
123.	Server Co-location	Server colocation is the process of deploying and hosting an organization-owned server within a managed service facility/environment. It enables an organization to deploy their servers within an existing data center or IT facility.
124.	Artificial Intelligence	Artificial intelligence is intelligence—perceiving, synthesizing, and inferring information—demonstrated by machines, as opposed to intelligence displayed by animals and humans.
125.	Machine Intelligence	Machine intelligence is advanced computing that enables a technology (a machine, device, or algorithm) to interact with its environment intelligently, meaning it can take actions to maximize its chance of successfully achieving its goals.
126.	Audit trail	An audit trail is a security-relevant chronological record, set of records, and/or destination and source of records that provide documentary evidence of the sequence of activities that have affected at any time a specific operation, procedure, event, or device
127.	Infrastructure-As-a service	Infrastructure as a service (IaaS) is a type of cloud computing service that offers essential compute,

Sl. No.	Term	Details
		storage, and networking resources on demand, on a pay-as-you-go basis.
128.	Platform-as-a-service	Platform as a service (PaaS) is a cloud computing model where a third-party provider delivers hardware and software tools to users over the internet.
129.	Software-as-a-service	Software as a service is a software licensing and delivery model in which software is licensed on a subscription basis and is centrally hosted. SaaS is also known as "on-demand software" and Web-based/Web-hosted software.
130.	Data mining	Data mining is the process of extracting and discovering patterns in large data sets involving methods at the intersection of machine learning, statistics, and database systems.
131.	Predictive Analytics	Predictive analytics encompasses a variety of statistical techniques from data mining, predictive modeling, and machine learning that analyze current and historical facts to make predictions about future or otherwise unknown events
132.	Prescriptive Analytics	Prescriptive analytics is a form of business analytics which suggests decision options for how to take advantage of a future opportunity or mitigate a future risk, and shows the implication of each decision option.
133.	Descriptive Analytics	Descriptive Analytics is the examination of data or content, usually manually performed, to answer the question “What happened?” (or What is happening?), characterized by traditional business intelligence (BI) and visualizations such as pie charts, bar charts, line graphs, tables, or generated narratives.

8 Appendix- State Digital Transformation Index (SDTI)

8.1 Introduction

Digital transformation has been changing the way we work, way we communicate, maintain economies, and keep government running. The World bank mentions that the concept of digital government represents a fundamental shift in the way governments around the world are embracing their mission. From setting measurable administrative goals to improving public service delivery, from making data-driven decisions to enacting evidence-based policies, from ensuring greater accountability and transparency within government to building greater public trust, governments are leveraging the power of information technologies in transformative ways.

Gartner mentions the key technologies to support digital transformation in the government to include,

- *Digital government technology platforms*, which are a set of cross-cutting, integrated, horizontal capabilities that coordinate government services across multiple domains, such as citizen experience, ecosystem, Internet of Things, and IT systems and analytics.
- *Digital citizen* identity to support online authentication in interactions with government, and increasingly across sectors and jurisdictions, overcoming past identity silos.
- *Hybrid cloud computing,* i.e., one or more public and private cloud services that operate as separate but integrated entities, offering a combination of cost optimization, agility and scalability of public cloud with control and compliance that are typical of the private cloud.
- *Data sharing* across the whole government, leveraging multiple data sources to create new value and achieve improved outcomes. This requires government departments to be willing to expose data and be able to analyze it through a systematic and scalable approach that can reuse data and innovate services.
- *Total experience* is a strategic approach designed to improve the engagement of citizens by providing them with modern tools across multiple channels and touchpoints that enhance overall experience, inclusion and equity. Moving away from a singular focus on customer experience improves the chances that governments can improve the quality of their services and capacity to deliver on their mission in the future.

Government of India has launched a flagship program named Digital India with a vision to transform India into a digitally empowered society and knowledge economy. Digital India encompasses key vision areas, Digital Infrastructure as a Core Utility to every citizen, Governance and Services On-Demand and Digital Empowerment of Citizens. Government of Assam also has a keen focus towards digitization of services provided to citizens, as well as digitization of all the government development programs. The State Digital Transformation Index (SDTI) is an objective effort to measure the digital maturity of the departments and implementing agencies.

This index is influenced and aligned to the Data Governance Quality Index (DGQI) being prepared by NITI Aayog to assess the preparedness of the various departments/ministries for usage of data in implementation of CS/CSS schemes of Government of India. The SDTI will be applicable to rank the various state departments and directorates providing public services and implementing development programs across the state using data for policy formulation and implementation.

8.2 Objectives and scope

a. **Objectives:** The aim of this exercise is to promote Digital economy in Assam. The digital economy as we know it, is the economic activity that results from billions of everyday online connections among people, businesses, government, devices, data, and processes. The backbone of the digital economy is hyperconnectivity which means growing interconnectedness of people, organisations, and machines that results from the Internet, mobile technology and newer technologies such as the internet of things (IoT).

This index has been developed with the following objectives:

- To measure the ability of the departments to integrate with the digital economy.
- To enhance the capabilities of the departments in evidence-based policy making and implementation monitoring.
- To enable review and assessment of preparedness of the data systems within the Departments on objective parameters of a standardized framework.

b. **Scope of this exercise**

This index has primarily been developed for State Government Departments. The assessment exercise can be undertaken by the nodal agency to ensure participation either from all the departments or a selected group of departments based on their scale and scope of work. Even within the participating departments, the nodal agency may decide to either include all the schemes and services of each of the department or a selected set of schemes/services based on the budget and nature of service or development schemes. Typically, it is prescribed that a pareto analysis should be conducted by the departments for the identifying the most critical schemes.

8.3 Approach and Methodology

a. **Overall Approach:** The key pillars of the assessment are identified as, Data Strategy, Data Systems, Data Outcomes. This theory of change will form the basis of the design of SDTI.

b. Indexing Methodology: Under the realm of the overall approach, the below six key themes have been identified under the data systems pillar which will have an impact on the SDTI.

- *Data Capture:* Data generation measures the ability of the respective departments to efficiently generate useful data in the course of their policy implementation. It covers areas related to the level of digitization, frequency and granularity of data generation, usage of multi-channel mode for capturing data, etc.

- *Data Quality*: Data Quality covers processes of scientifically and statistically evaluating data in order to determine whether they meet quality benchmarks. The key areas covered under this theme relate to profiling of data, data quality assessment processes (for e.g., data pipeline design, well defined data schema, capturing metadata, etc.), data cleaning, use of latest technologies in the process.

- *Data Interoperability:* Data Interoperability is defined as the ability of systems and services that create, exchange and consume data to have clear, shared expectations for the contents, context, and meaning of that data. Simply put, it is when data in diverse formats and from different sources can be unified and used together for policy making in Government context.

- *Data Analysis and Reporting*: This theme covers if the collected data is being analyzed and used for evidence creation and decision making. Given the present context, it measures the capability of departments are undertaking basic cross-sectional analyses only or regression and predictive analysis as well. The use of dashboards for visualization of data is also checked to ensure that information is disseminated in a user-friendly manner. It also assesses if other social media platforms are also being increasingly used for information dissemination and whether websites have features to support multilingual interfaces and are GIGW compliant. This also measures whether the evidence created are being used in reporting to senior officers, and actions being taken for formulation or course-correction.

- *Data Security:* Data security assessment and audit is an independent process in itself, and requires organizational focus. The current theme is to assess the compliance of the departments to Data security guidelines provided as a part of the Operational guidelines of the State Data policy.

- *Capacity Building*: The sustainability of the evidence-based policy making requires inculcation of specialized skills in data science, statistics as well as information technology to be built within the organization. This theme assesses the HR Capacity of the department to promote a data driven decision making across the organization.

The various themes have been divided into subsections with each subsection given weightages. The weightage has been given based on the importance each of these themes play in implementation of operational guidelines of the Assam State Data Policy, 2022, and importance of these themes in institutionalizing evidence-based policy formulation and evaluation. SDTI may be reviewed from time-to-time based on the various national and international indices to measure maturing of Data Systems.

The subsections of the themes and the corresponding weightages in the index has been detailed out as below:

Theme	Theme Weightage	Sub-section
Data Capture	15	Granularity and Digitization
		Frequency of Update
Data Quality	20	Data quality protocols/use of metadata
		Treatment of Personal/Sensitive and Publishing Open data
Data Interoperability	30	Adherence to best practices for building data systems
		Data sharing with other departments.
Data Analysis and Reporting	15	Various techniques and practices of data analytics and dashboards
Data Security	10	Best practices of data security being adopted
Capacity Building	10	Building skills through recruitment and training.

c. How the assessment will be carried out?

This assessment will be carried out on frequent intervals to measure the readiness of the state departments. The department would be scored and ranked on the basis of this assessment. This assessment would be initiated out by the Center of Data Management within a span of 15 days after the notification. The assessment report would be submitted to all the departments, as well as the Chief Minister's Office.

- *Baseline assessment:* All the departments would be baselined within two months of publishing of this index.
- *Periodic Yearly assessment:* The various departments would be notified on the Yearly assessment dates in advance. And departments need to comply to procedures of assessment survey to be carried out by the nodal department.

8.4 Annexure 1: Indicative Draft of Assessment Questionnaire

Part A- General Information

1	Department	<abc>
2	Name of the schemes	a) xxx b) yyy
3	Name of the services (G2C, G2B, G2G)	a) zzz b) aaa

Part B- Scheme/Service Information

A. Data Capture: Granularity and Digitization

1.	At what Granularity the data is generated for this scheme/service	Paper	Digital	
	a) State Level	Yes □	Yes □	
	b) District	Yes □	Yes □	N/A □
	c) Block	Yes □	Yes □	N/A □
	d) Gram Panchayat	Yes □	Yes □	N/A □
	e) Village	Yes □	Yes □	N/A □
	f) Individual	Yes □	Yes □	N/A □
	g) Assisted	Yes □	Yes □	N/A □
2.	At what Granularity the data is digitized for this scheme/service	Paper	Digital	
	a) State Level	Yes □	Yes □	
	b) District	Yes □	Yes □	N/A □
	c) Block	Yes □	Yes □	N/A □
	d) Gram Panchayat	Yes □	Yes □	N/A □
	e) Village	Yes □	Yes □	N/A □
	f) Individual	Yes □	Yes □	N/A □
3.	Usage of mobile phones/ tablets	Yes □	Yes □	N/A □
	If 'Yes', what are the purposes for which mobile phones are used			
	a) Outreach – IEC campaigns	Yes □	No □	N/A □
	b) Applications/data collections	Yes □	No □	N/A □
	c) Feedback	Yes □	No □	N/A □
	d) Telephonic survey (manual/IVR)	Yes □	No □	N/A □
	e) Geotagged photograph uploading	Yes □	No □	N/A □
	f) Geo fenced data generation	Yes □	No □	N/A □
	g) Location and GPS data	Yes □	No □	N/A □

	h) Multimedia data – voice, video, images as evidence	Yes □	No □	N/A □
4.	Record Keeping			
	Is e-office being used for record keeping? If yes,	Yes □	No □	
	a) For tracking Physical files only	Yes □	No □	
	b) For sharing of E-files within the department/ directorate/ office	Yes □	No □	
	c) For sharing the E-files across the departments	Yes □	No □	N/A □

B. Data Capture: Frequency

1.	At what frequency is data generated	Paper	Digital
	a) Real time (transaction data)	Yes □	Yes □
	b) Daily	Yes □	Yes □
	c) Weekly/Fortnightly	Yes □	Yes □
	d) Monthly	Yes □	Yes □
	e) Quarterly	Yes □	Yes □
	f) Half-yearly	Yes □	Yes □
	g) Yearly	Yes □	Yes □

C. Data Quality – Protocols for Data quality and data definition

1.	Does the scheme/surveys use protocols to use data quality	Yes □	No □
	a) Rigorous data profiling and control of incoming data	Yes □	
	b) Data pipeline design to avoid duplicate data	Yes □	
	c) Accurate gathering of data requirements (well-defined data schema)	Yes □	
	d) Enforcement of data integrity	Yes □	
	e) Integration of data lineage traceability into the data pipeline	Yes □	
	f) Creation and use of metadata, and data dictionary	Yes □	

	g) Dedicated data quality teams	Yes □	
2.	Usage of Data standardization	Yes □	No □
	a) Local Government directory (LGD) for addresses	Yes □	

D. Data Quality – Treatment of Personal/Sensitive and Publishing Open data

1.	Does the scheme/surveys use protocols for handling personal data	Yes □	No □
	a) Capture of consent for usage of personal data based on prevailing norms for data protection and privacy	Yes □	
	b) Adequate security controls for restricting access to personal data in the MIS/Data systems while data at transit as well as data at rest	Yes □	
	c) Usage of anonymization and aggregation techniques for de-identification of data for doing data analysis	Yes □	
	d) Redressal mechanism for grievances regarding personal data within predefined timeline as necessitated by prevailing norms for data protection and privacy.	Yes □	
	e) Controls for handling of sensitive data/ adherence to national and state laws for processing of sensitive data	Yes □	
2.	Does the scheme/surveys use publishes data regularly into OGD portal and State Data analytics portal?	Yes □	No □
3.	Data provenance		
	a) Usage of e-sign for documents, wherever required	Yes □	
	b) Maintenance of audit trail	Yes □	

E. Data Interoperability – Adherence to best practices for building data systems.

1.	Do the data systems used by the Department for implementation of schemes/delivering services adherence to best practices of building data systems?		
	a) Usage of data registries for building single source of truth	Yes □	

	b) Usage of interoperability standards like OpenAPI by default	Yes □	
	c) Federated access for the registries is available for other state departments	Yes □	
	d) Seeding of digital ID across the department's registries for implementation of social protection programs	Yes □	
	e) Implementation of virtual credentialing	Yes □	
2.	Data System characteristics		
	a) Is it Open – Source?	Yes □	
	b) Is the application for the Data system being developed by an out-sourced agency? If yes, does the department own the IPR of the application?	Yes □	
	c) Can the code of the application be reused by the State for another requirement?	Yes □	
3.	Does the scheme/surveys use data systems hosted at Cloud?	Yes □	No □
	If yes, what services are being used		
	a) Co-location of the servers	Yes □	
	b) Private Cloud at SDC/NDC	Yes □	
	c) Government Community Cloud (GCC)/Virtual private Cloud at MeitY empanelled CSP	Yes □	
	d) Infrastructure-As-a service	Yes □	
	e) Platform-as-a-service	Yes □	N/A □
	f) Software-as-a-service	Yes □	N/A □
4.	Number of users of the Data system for the scheme implementation/service delivery		
	a) Users at State level	Yes □	>100 □
	b) Users at the district level	Yes □	>50 □
	c) Users at the sub-district/Block level	Yes □	>20 □
	d) Users at village level	Yes □	>10 □
5.	Use of advanced technologies like AI, Blockchain, IoT ,etc		
	a) Usage of Artificial Intelligence (AI)	Yes □	N/A □

	b) Usage of Blockchain	Yes □	N/A □
	c) Usage of Internet of Things (IoT)	Yes □	N/A □
	d) Usage of Big Data analytics	Yes □	N/A □

F. Data Interoperability – Data sharing with other departments.

1.	Do the Departments' data systems used for implementation of schemes/delivering services adhere to the Operational Guidelines for the Assam State Data Policy 2022?		
	a) Establishment of Implementation Authorities as mandated by these guidelines	Yes □	No □
	b) Identification and Classification of datasets as mandated by the guidelines	Yes □	No □
	c) Notification of data standards for field level data elements	Yes □	No □
	d) Compliance to Data Governance Framework as mandated by these guidelines	Yes □	No □
	e) Compliance to Data Storage Requirements for dataset received as per these guidelines	Yes □	No □
	f) Compliance to Data Exchange Framework as per these guidelines	Yes □	No □
	g) Notification of Departmental registries as single source of truth as per the guidelines	Yes □	No □
	h) Notification of Business Continuity/IT recovery plan. If yes, then		
	• State level	Yes □	No □
	• Department level	Yes □	No □
	• District level	Yes □	No □
	i) Notification and compliance to Metadata standards in capturing data as per these guidelines	Yes □	No □
	j) Establishment of Training modules and workshops	Yes □	No □
	k) Archiving datasets in Assam State Datalake for Inter-departmental analysis	Yes □	No □

2.	Do the Departments' data systems adhere to the Operational guidelines for sharing of Data?		
	a) As open data in the State OGD portal	Yes □	No □
	b) As permissioned-access data with the requesting organizations, and the count of the datasets shared are:		
	• Less than 5	Yes □	No □
	• Between 5 to 10	Yes □	No □
	• Between 10 to 50	Yes □	No □
	• More than 50	Yes □	No □

G. Data Analysis and Reporting.

1.	Do the scheme/service delivery uses data analysis techniques	Yes □	No □	
2.	If yes, what are the methods used			
	a) Exploratory data analysis	Yes □	No □	N/A □
	b) Modelling and Algorithms	Yes □	No □	N/A □
	c) Correlation	Yes □	No □	N/A □
	d) Causation	Yes □	No □	N/A □
	e) Regression Analysis	Yes □	No □	N/A □
	f) Predictive	Yes □	No □	N/A □
	g) Data mining	Yes □	No □	N/A □
3.	Does the scheme/service delivery use dashboards	Yes □	No □	
4.	If yes, what are the purposes for which dashboards are being used			
	a) Visual presentation of performance measures	Yes □		
	b) Identifying pre-empt trends	Yes □		
	c) Measure efficiencies /inefficiencies	Yes □		
	d) Generate detailed reports showing new trends	Yes □		
	e) Make more informed decisions based on collected business intelligence	Yes □		
	f) User friendly one stop access to multiple reports	Yes □		

	g) Gain total visibility of all systems instantly	Yes □		
	h) Quick identification of data outliers and correlations	Yes □		
5.	Use of other (external) data sources for data analysis	Yes □	No □	
	a) Remote sensing /Geo- spatial data	Yes □	No □	N/A □
	b) IoT data	Yes □	No □	N/A □
	c) Social media data	Yes □	No □	N/A □
	d) Private sector generated data	Yes □	No □	N/A □

H. Data Security – Best practices of data security being adopted

1.	Data Security measures being taken by the Data Systems having government data?		
	a) Adherence to good programming practices to avoid OWASP top 10 vulnerabilities?	Yes □	No □
	b) Regular security audit by a CERT empaneled auditor?	Yes □	No □
	c) Proactive monitoring against DDoS attacks, usage of subnetting, firewalls	Yes □	No □
	d) Implementation of IAM and multi-factor authentication	Yes □	No □

I. Capacity Building

1.	Measures being taken for Departmental capacity building		
	a) Recruitment of staff with technical skillset on government payroll.	Yes □	No □
	b) Recruitment of staff with Data Analytics skillset on government payroll.	Yes □	No □
	c) Regular training programs for the staff	Yes □	No □
	d) Inclusion of Data analysis and reporting as a parameter in the Annual performance report of the staff	Yes □	No □

8.5 Annexure 2: Scoring mechanism

Sl.No.	Block	Scoring Logic
1	Data Generation: Granularity and Digitization	Scoring is to be done primarily for the Digital Part Only. If the Scheme/Department is collecting Paper-based information at any granularity at State or Sub-state level, give '0'. Scoring for Usage of mobile phones/ tablets must be answered in 'Yes', unless it is 'not applicable' for receiving marks, otherwise no marks are to be awarded.
2	Data Capture: Frequency	Scoring for frequency would also be done for Digital part only, and any response which is paper based would be given '0'.
3.	Data Quality – Protocols for Data quality and data definition	Scoring would be done in the following manner: ⇒ If response to Q1 a), b), c), e) are 'Yes', then 1 mark would be awarded. ⇒ If response to Q1 d), f), g) are 'Yes', then 2 mark would be awarded. ⇒ If response to Q2 is 'Yes, then 1 mark would be awarded.
4.	Data Quality – Treatment of Personal/Sensitive and Publishing Open data	Scoring would be done in the following manner: ⇒ If response to Q 1. a), to 1.e) are 'Yes', then 2 mark would be awarded. ⇒ If response to Q 2 are 'Yes', then 1 mark would be awarded. ⇒ If response to Q 3 a) and b) are 'Yes', then 2 mark would be awarded.
5.	Data Interoperability – Adherence to best practices for building data systems.	Scoring would be done in the following manner: ⇒ If response to Q 1. a), to 1.e) are 'Yes', then 3 marks would be awarded. ⇒ If response to Q 2 a) are 'Yes', then 1 mark would be awarded. ⇒ If response to Q2 b) and c) are 'Yes', then 2 marks would be awarded. ⇒ If response to Q 3 a) and b) are 'Yes', then 2 mark would be awarded. ⇒ If response to Q 3 c), d) and e) are 'Yes', then 1 mark would be awarded. ⇒ If response to Q 4 a) to d) are 'Yes', then 1 mark would be awarded, if the response to the supplementary response on number of user is also selected, then 2 mark would be awarded. ⇒ If response to Q 5 c), d) and e) are 'Yes', then 1 mark would be awarded.

6.	Data Interoperability – Data sharing with other departments.	Scoring would be done in the following manner: ⇒ If response to Q 1. a), to c) are 'Yes', then 2 marks would be awarded.
7.	Data Analysis and Reporting.	Scoring would be done in the following manner: ⇒ If response to Q 1. are 'Yes', then 1 marks would be awarded. ⇒ If response to Q 2 a) to g). are 'Yes', then 1 marks each would be awarded. ⇒ If response to Q 3. are 'Yes', then 1 marks would be awarded. ⇒ If response to Q 4 a) to h). are 'Yes', then 1 marks would be awarded. ⇒ If response to Q 5 a) to d). are 'Yes', then 1 marks would be awarded.
8.	Data Security – Best practices of data security being adopted	Scoring would be done in the following manner: ⇒ If response to Q 1. a) to f) are 'Yes', then 2 marks would be awarded. ⇒ If response to Q 1. g) to k) are 'Yes' including the sub-questions, then 1 marks would be awarded. ⇒ If response to Q 2. a) to b) are 'Yes' including the sub-questions, then 2 marks would be awarded.
9.	Capacity Building	Scoring would be done in the following manner: ⇒ If response to Q 1. a) to c) are 'Yes', then 1 marks would be awarded. ⇒ If response to Q 1. d) are 'Yes', then 1 marks would be awarded.

ANURAG GOEL,
Principal Secretary to the Government of Assam,
Information Technology Department.

Guwahati : Printed and Published by the Director, Directorate of Printing & Stationery, Assam, Guwahati-21.
Extraordinary Gazette No. 407 - 10 + 10 - 29 - 04 - 2023. (visit at– dpns.assam.gov.in)